理學博士學位論文

Structural Studies of Three Cell Shape Determining Proteins from *Helicobacter pylori*: Csd1, Csd2, and Csd3

2016年 8月

서울大學校 大學院

生物物理 및 化學生物學科 生化學專攻

안 두 리

Structural Studies of Three Cell Shape Determining Proteins from *Helicobacter pylori*: Csd1, Csd2, and Csd3

Thesis by

Doo Ri An

Professor: Se Won Suh

A Thesis Submitted to the Graduate Faculty of Seoul National University in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2016

Abstract

Helicobacter pylori is associated with various gastrointestinal diseases, including gastric cancer. Its colonization of the human gastric mucosa requires high motility, which depends on the helical cell shape. In *H. pylori*, cell shape-determining genes (*csd1*, *csd2*, *csd3*/*hdpA*, *ccmA*, *csd4*, *csd5*, and *csd6*) play key roles in determining the cell shape by alteration of cross-linking or by trimming of peptidoglycan stem peptides. Among them, Csd1, Csd2, and Csd3/HdpA belong to the same M23B metallopeptidase family and may act as D,D-endopeptidases to cleave the cross-linking (D-Ala^4-*m*DAP^3 bond) of cross-linked dimer muropeptides. Csd3 is a bi-functional enzyme, and it functions also as the D,D-carboxypeptidase to cleave the D-Ala^4-D-Ala^5 bond of the muramyl pentapeptide. To better understand the role of Csd proteins in *H. pylori*, I have determined the crystal structures of Csd1 (HP1543 in 26695 strain), Csd2 (HP1544), and Csd3 (HP0506). *H. pylori* Csd3 exists as monomers in solution. The crystal structure of Csd3 revealed that the Csd3 monomer consists of three domains: domain 1, domain 2, and C-terminal LytM domain. The Csd3 LytM domain contains the catalytic active site with a Zn^{2+} ion, like similar domains of other M23 metallopeptidases. However, the active site in the Csd3 LytM domain is

blocked by domain 1, resulting in a latent and inactive state. *H. pylori* Csd2 alone exists in monomer-dimer equilibrium and forms a stable heterodimer with *H. pylori* Csd1 in solution. The crystal structures of $Csd2_{121–308}$ homodimer and $Csd1_{125–312}$-$Csd2_{121–308}$ heterodimer revealed that overall structures of $Csd1_{125–312}$ and $Csd2_{121–308}$ monomers are similar to each other, consisting of a helical domain and a LytM domain. The helical domains of both Csd1 and Csd2 play a key role in the formation of homodimers or heterodimers. LytM domains of Csd1 and Csd2 share the same overall fold but a functionally significant difference exists in their active sites. The Csd1 LytM domain contains a catalytic site with a Zn^{2+} ion, which is coordinated by three conserved ligands and two water molecules, whereas the Csd2 LytM domain has incomplete metal ligands and no metal ion is bound. Although Csd3 and Csd1/Csd2 proteins are LytM homologs in *H. pylori*, they show distinct features in their domain organization and LytM domains. Structural knowledge of these Csd proteins sheds light on the events that regulate the cell wall in *H. pylori*.

Keywords: *Helicobacter pylori* / cell shape determinant / Csd1 / Csd2 / Csd3 / HP1543 / HP1544 / HP0506 / peptidoglycan / D,D-endopeptidase / D,D-carboxypeptide / M23B family metallopeptidase / LytM domain

Table of Contents

Abstract · · ·

Chapter 1. Crystal Structure of Csd3 from *Helicobacter pylori*, a Cell Shape-Determining Metallopeptidase

1.1 Introduction · · ·

Chapter 2. Structural Basis of the Heterodimer Formation between Cell Shape-Determining Proteins Csd1 and Csd2 from *Helicobacter pylori*

2.1 Introduction · · ·

2.3.7 Discussion · · ·

List of Tables

Chapter 1.

Chapter 2.

Table 2-1. Construct design and protein expression for *H. pylori* Csd2 · · · 91

Table 2-2. Construct design and protein expression for *H. pylori* Csd1 · · · 92

Table 2-3. Construct design and protein expression for Csd1-Csd2 complex · · · · · · · · · · · · · · · · · ·

List of Figures

Chapter 1.

Figure 1-1. *Helicobacter pylori* infection · 3

Figure 1-2. Strategies for *H. pylori* infection · 5

Figure 1-3. Csd proteins function as peptidoglycan hydrolases in *H. pylori* · · · · · · · · · · · · · · · · · · ·

Figure 1-16. Conserved residues in the active site of M23B metallopeptidases · · ·

Chapter 2.

Figure 2-1. Chemical structure of cross-linked peptidoglycan in Gram-negative bacteria · 78

Figure 2-2. Cleavage specificities of peptidoglycan hydrolases in *H. pylori* · · · · · · · · · · · · · · · · · · ·

Abbreviations

Csd	cell shape determinant
NAG	*N*-acetylglucosamine
NAM	*N*-acetylmuramic acid
D,L-CPase	D,L-carboxypeptidase
D,D-EPase	D,D-endopeptidase
D,D-CPase	D,D-carboxypeptidase
L,D-CPase	L,D-carboxypeptidase
Tris	tris(hydroxymethyl)aminomethane
PEG	polyethylene glycol
R.m.s	root-mean-square
SAD	single-wavelength anomalous diffraction
SEC-MALS	size-exclusion chromatography with multi-angle static light scattering
HPLC	high-performance liquid chromatography
PDB	Protein Data Bank

Chapter 1

Crystal Structure of Csd3 from *Helicobacter pylori*, a Cell Shape-Determining Metallopeptidase

1.1. Introduction

Helicobacter pylori is a helical-shaped Gram-negative bacterium that colonizes the human stomach (Figure 1-1). It colonizes approximately half of the world's population and its infection of the gastric mucosa has been associated with various diseases of the upper gastrointestinal tract, such as chronic gastritis, peptic ulcer, mucosa-associated lymphoid tissue (MALT) lymphoma, and gastric adenocarcinoma (Roesler et al., 2014; Kusters et al., 2006). *H. pylori* was classified as a group I carcinogen by The International Agency for Research on Cancer and is regarded as a primary factor for gastric cancer development (Roesler et al., 2014). *H. pylori* infection has also been implicated with some extra-digestive diseases (Roubaud Baudron et al., 2013). Relatively effective treatment regimens are available for *H. pylori* infection that typically consists of a proton pump inhibitor such as omeprazole and the antibiotics clarithromycin and amoxicillin (or

metronidazole). However, increasing antibiotic resistance requires new therapies and discovery of new antibiotics (Malfertheiner et al., 2012).

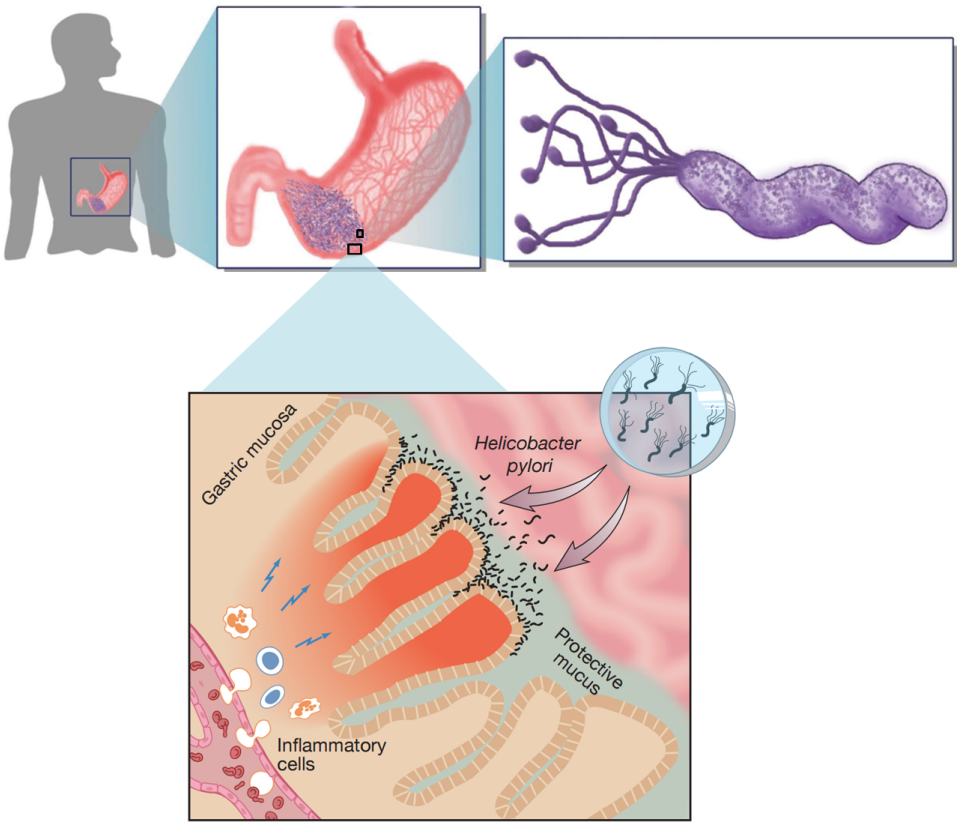


Figure 1-1. *Helicobacter pylori* infection (Ayala et al., 2014 and http://www.nobelprize.org/)

Helicobacter pylori infects the lower part of the stomach. Infection of the stomach by *H. pylori* persists lifelong and causes gastric inflammation (gastritis). Chronic gastritis is likely to underlie *H. pylori*-related diseases.

High motility of *H. pylori* is important for its colonization of the human stomach and its survival in its preferred niche, the gastric mucosa (Ottemann and Lowenthal, 2002; Schreiber et al., 2004; Lertsethtakarn et al., 2011). The helical cell shape of *H. pylori* is believed to facilitate efficient colonization of the viscous epithelial mucus layer via a cork-screwing mechanism (Berg and Turner, 1979; Hazell et al., 1986; Worku et al., 1999) (Figure 1-2). Mutants of *H. pylori* with altered cell shapes exhibited attenuated colonization (Sycuro et al., 2010; Wyckoff et al., 2012).

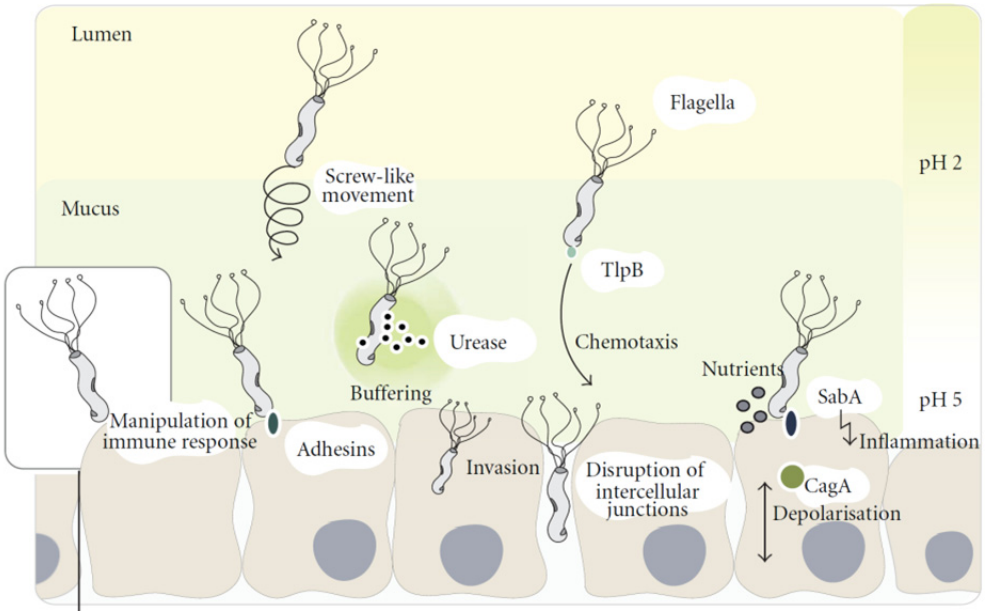


Figure 1-2. Strategies for *H. pylori* infection (Bauer and Meyer, 2011)

H. pylori has evolved numerous strategies to colonization and persistence in the hostile environment of the human stomach.

The peptidoglycan layer of a bacterial cell wall plays a key role not only in protecting cells against the intracellular turgor pressure but also in maintaining the proper cell shape (Scheffers and Pinho, 2005; Vollmer and Bertsche, 2008). It is made of linear polysaccharide chains that consist of alternating β-1,4-linked *N*-acetylglucosamine-*N*-acetylmuramic acid (NAG-NAM) disaccharide units, with a pentapeptide covalently linked to NAM (Vollmer et al., 2008a). The pentapeptide chains are either 4→3 or 3→3 cross-linked to different extents depending on bacterial species. In *H. pylori*, the pentapeptide consists of L-Ala^{1}-γ-D-Glu^{2}-*m*DAP^{3}-D-Ala^{4}-D-Ala^{5} (or -Gly^{5}), where mDAP refers to meso-2,6-diaminopimelate. The neighboring peptides are cross-linked exclusively by the 4→3 linkage between the α-carboxylic group of D-Ala^{4} on one strand (the donor peptide) and the ε-amino group of *m*DAP^{3} on another strand (the acceptor peptide) (Weidel and Pelzer, 1964; Costa et al., 1999) to form a mesh-like peptidoglycan (murein) sacculus (Meroueh et al., 2006). In many bacteria, the peptidoglycan layer is remodeled by a number of peptidoglycan hydrolases that function in the peptidoglycan maturation, regulation of cell wall growth, cell division, peptidoglycan turnover and recycling, cell lysis and the release of peptidoglycan fragments for host-pathogen interactions (Vollmer et al., 2008b; Vollmer, 2012; Lee and Huang, 2013; Heijenoort, 2011).

In *H. pylori*, a series of genes have been identified to be required for determining the helical cell shape: *csd1*, *csd2*, *csd3/hdpA*, *ccmA*, *csd4*, *csd5*, and *csd6* (named cell shape-determinant genes or *csd*'s and curved cell morphology gene or *ccmA*) (Sycuro et al., 2012; Sycuro et al., 2010; Sycuro et al., 2013; Bonis et al., 2010). Deletion mutants for each single gene resulted in the impaired morphology (curved rod in *Δcsd1*, *Δcsd2*, and *ΔccmA*; irregular c-shaped or stocky branched morphology in *Δcsd3*; straight rod in *Δcsd4*, *Δcsd5*, and *Δcsd6*), and showed changes in peptidoglycan compositions. They play important roles in determining the helical cell shape by relaxation of peptidoglycan crosslinking or by trimming of peptidoglycan muropeptides (Figure 1-3). *H. pylori* Csd6 functions as the L,D-carboxypeptidase (L,D-CPase) that cleaves the *m*DAP^3-D-Ala^4 bond of the uncrosslinked muramyl tetrapeptide (muramyl-L-Ala^1-γ-D-Glu^2-*m*DAP^3-D-Ala^4) to produce the muramyl tripeptide (muramyl-L-Ala^1-γ-D-Glu^2-*m*DAP^3) and D-Ala (Sycuro et al., 2013; Kim et al., 2015). *H. pylori* Csd4 is a Zn^{2+}-dependent D,L-carboxypeptidase (D,L-CPase) that cleaves the γ-D-Glu^2-*m*DAP^3 bond of the uncrosslinked muramyl tripeptide (muramyl-L-Ala^1-γ-D-Glu^2-*m*DAP^3) to produce the muramyl dipeptide (muramyl- L-Ala^1-γ-D-Glu^2) and *m*DAP (Sycuro *et al.*, 2012; Kim et al., 2014; Chan et al., 2015). *H. pylori* Csd5 has no known enzymatic activity, but interacts physically with

both *H. pylori* Csd4 and a dipeptide product of the reaction catalyzed by Csd4 (Kim et al., 2014). It has been suggested that *H. pylori* Csd5 may play a regulatory role in modulating the Csd4 function (Kim et al., 2014). *H. pylori* CcmA does not encode a putative enzyme and CcmA-like bactofilins form cytoplasmic filaments that bind a peptidoglycan synthesis enzyme for the localized activity (Koch et al., 2011; Kühn et al., 2010). It may exert its function by interacting with the Csd proteins, providing stabilization and/or localization (Sycuro *et al.*, 2010).

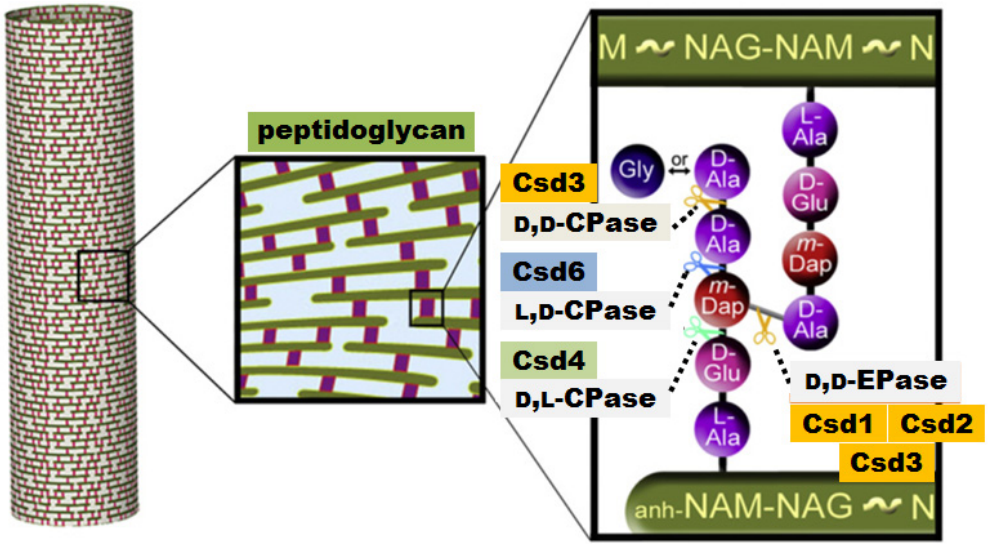


Figure 1-3. Csd proteins function as peptidoglycan hydrolases in *H. pylori* (modified from Sycuro et al., 2010)

Schematic illustration of cell wall (left) and peptidoglycan structure (middle) that is composed of glycan strands (green) interconnected by peptide crosslinks (purple). Tetra-pentapeptide muropeptide is shown with the peptide crosslink highlighted in grey (right). Peptide bonds marked with the scissors can be hydrolyzed by Csd proteins.

H. pylori Csd3, as well as Csd1 and Csd2, belongs to a M23B metallopeptidase family (Sycuro et al., 2010). Deletion of *csd1*, *csd2*, and *csd3* genes eliminated an D,D-endopeptidase (D,D-EPase) activity, which cleaves the D-Ala^4-*m*DAP^3 peptide bond (4→3 linkage) in crosslinked dimers of muramyl tetra- and penta-peptides (muramyl-L-Ala^1-γ-D-Glu^2-mDAP^3-D-Ala^4 and muramyl-L-Ala^1-γ-D-Glu^2-mDAP^3-D-Ala^4-D-Ala^5, respectively) (Sycuro et al., 2010). Interestingly, the muropeptide composition for the deletion mutant of the *csd3* gene (*Δcsd3*) indicated that Csd3 has an additional D,D-carboxypeptidase (D,D-CPase) activity to cleave the D-Ala^4-D-Ala^5 bond of the muramyl pentapeptide to produce the muramyl tetrapeptide and D-Ala (Bonis et al., 2010; Sycuro et al., 2010). In accordance with this observation, *Δcsd3* showed irregular c-shaped or stocky branched cells, which are distinct from the curved rod morphology possessed by *Δcsd1* and *Δcsd2* (Bonis et al., 2010; Sycuro et al., 2010) (Figure 1-4). Because of the bi-functional peptidase activities, *H. pylori* Csd3 is also known as HdpA (*Helicobacter* D,D-peptidase A). Inactivation of Csd3 by H259A mutation, which is predicted to affect metal coordination in the active site, also resulted in the same degree of morphological abnormality as in *Δcsd3* (Bonis et al., 2010). The bi-functional peptidase activities of Csd3,

together with this observation, makes Csd3 an important regulator of *H. pylori* morphology.

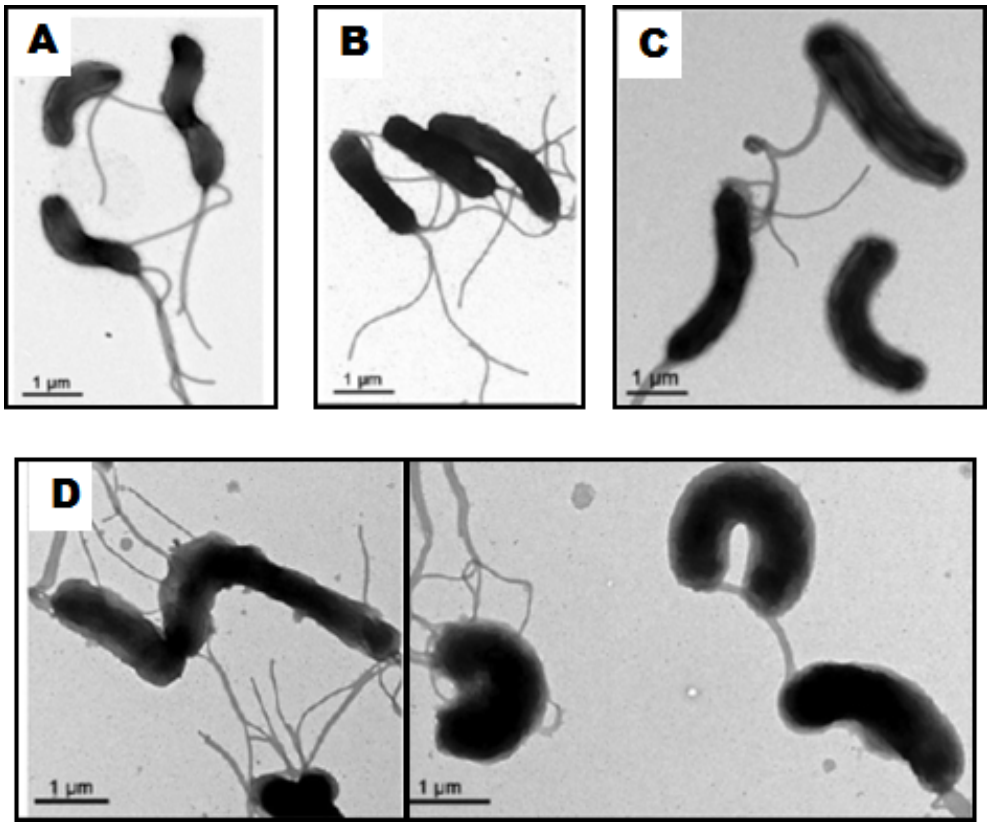


Figure 1-4. Morphologies of Csd1, Csd2, and Csd3 mutants

Transmission electron microscopy (TEM) images of wild-type (A), *Δcsd1* (B), *Δcsd2* (C), and *Δcsd3* (D).

To provide a structural framework for a better understanding of the molecular function of *H. pylori* Csd3, I have determined the crystal structure of N-terminally truncated Csd3 encompassing residues 42–403 ($Csd3_{\Delta 41}$). It consists of three domains: domain 1 (residues Glu42–Ile124), domain 2 (residues Ile125–Gly228 and Ala360–Phe403), and the C-terminal LytM domain (residues Phe229–Thr359). The Csd3 domain 1 and the core of domain 2 (residues Ile125–Gly228) share a common fold, despite a very low level of sequence identity. The LytM domain of Csd3 is structurally similar to the corresponding domains of other MEROPS M23 family metallopeptidases. Substrate binding to the active site of the LytM domain is blocked by domain 1 in this structure, suggesting that domain 1 is the inhibitory domain and that this Csd3 structure is in the latent state. The core of domain 2 is held stably against the LytM domain by the C-terminal extended tail region that protrudes from the LytM domain. This work could serve as the foundation in discovery of novel inhibitors that would prove helpful in fighting infections by the major human pathogen *H. pylori*.

1.2. Material and methods

1.2.1. Cloning, expression and purification

The *csd3* gene (*hp0506* from *H. pylori* 26695 strain) encoding the N-terminally truncated form (residues 42–403; $Csd3_{\Delta 41}$) was PCR-amplified and cloned into the expression vector pET-21a(+) (Novagen) using the NdeI and XhoI restriction enzyme sites. This construct contains a hexahistidine-containing tag (LEHHHHHH) at its C-terminus. The recombinant protein was overexpressed in *Escherichia coli* Rosetta 2(DE3)pLysS cells. The cells were grown at 37°C in Terrific Broth culture medium. Protein expression was induced by 0.5 mM isopropyl β-D-thiogalactopyranoside and the cells were incubated for additional 15 h at 30°C. The cells were lysed by sonication in ice-cold buffer A [20 mM Tris-HCl at pH 7.9, 500 mM sodium chloride, 50 mM imidazole, and 10% (*v/v*) glycerol], which was supplemented with 1 mM phenylmethylsulfonyl fluoride, 60 mM ammonium chloride, and 15 mM magnesium acetate. The lysate was centrifuged to discard the cell debris. The supernatant was applied to an affinity chromatography column of HiTrap Chelating HP (GE Healthcare), which was previously equilibrated with buffer A. The column was eluted with

buffer B [20 mM Tris-HCl at pH 7.9, 500 mM NaCl, 500 mM imidazole, and 10% (*v/v*) glycerol], with the $Csd3_{\Delta 41}$ protein being eluted at 125–150 mM imidazole concentration. The eluted protein was further purified by size-exclusion chromatography (gel filtration) using HiLoad 16/60 Superdex 200 prep-grade column (GE Healthcare), which was previously equilibrated with buffer C (20 mM Tris-HCl at pH 7.9 and 400 mM sodium chloride) or buffer D (20 mM Tris-HCl at pH 7.9, 200 mM sodium chloride and 0.1 mM zinc chloride), yielding two different batches of Csd3. Fractions containing the recombinant $Csd3_{\Delta 41}$ protein were pooled and concentrated to 10 mg ml^{-1} for crystallization using an Amicon Ultra-15 Centrifugal Filter Unit (Millipore).

1.2.2. Crystallization

Crystals were grown at 23°C by the sitting-drop vapor diffusion method using the Mosquito robotic system (TTP Labtech). Initial crystallization conditions were searched using Screen I, II, PEG/Ion, Index, Membfac, SaltRX, Natrix I, II, ScreenCryo (Hampton Research), and Wizard I, II, III, IV screening solutions (Rigaku Reagents). Two types of native crystals (Form 1 and Form 2) were produced under different crystallization

conditions. For Form 1 crystals, each sitting drop was prepared by mixing equal volumes (0.4 μl each) of the purified $Csd3_{\Delta 41}$ protein in buffer C and the reservoir solution [160 mM ammonium sulfate, 80 mM sodium acetate at pH 4.6, 20% (*v/v*) PEG 4000, and 20% (*v/v*) glycerol] and was equilibrated against 100 μl of the reservoir solution. Elongated rectangular crystals grew up to approximate dimensions of 0.3 mm × 0.2 mm × 0.3 mm within a few days (Figure 1-5). For Form 2 crystals, the purified $Csd3_{\Delta 41}$ protein purified in buffer D was pre-incubated with buffer D supplemented with 1 mM zinc chloride under ice for 30 min prior to crystallization setup. Sitting drops were prepared by mixing equal volumes (0.3 μl each) of the purified $Csd3_{\Delta 41}$ protein in buffer C and the reservoir solution [200 mM di-ammonium hydrogen phosphate at pH 7.9 and 20% (*v/v*) PEG 3350]. Hexagonal bipyramidal crystals grew up to approximate dimensions of 0.1 mm × 0.1 mm × 0.2 mm within a few days (Figure 1-5).

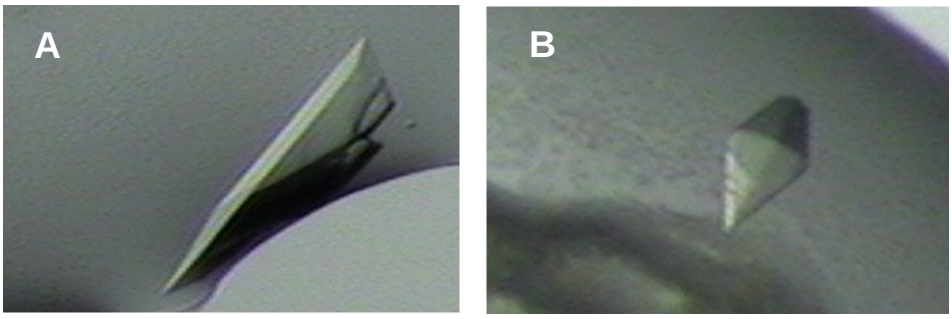


Figure 1-5. **Two native crystals of *H. pylori* Csd3$_{\Delta 41}$**

(A) Form 1 crystal with approximate dimensions of 0.3 mm × 0.2 mm × 0.3 mm. (B) Form 2 crystal with approximate dimensions of 0.1 mm × 0.1 mm × 0.2 mm.

1.2.3. X-ray data collection and structure determination

Native X-ray diffraction data were collected at the beamline BL-5A of Photon Factory, Japan using the ADSC Q315 CCD detector and at the beamline BL44XU of SPring-8, Japan using the DIP6040 imaging plate/CCD hybrid detector (Table 1-1). Form 1 crystals were flash-frozen in a nitrogen gas stream at 100 K. They diffracted to 2.0 Å and belong to the space group $P2_12_12_1$, with unit cell parameters of a = 62.6 Å, b = 112.1 Å, and c = 112.9 Å. Assuming the presence of two Csd3 monomers in the asymmetric unit, the Matthew's coefficient and the solvent content are 2.33 $Å^3$ Da^{-1} and 47.2 %, respectively. Form 2 crystals were soaked for several seconds in a cryoprotectant solution, which was the reservoir solution supplemented with 25% (*v/v*) glycerol, and were flash-frozen in a nitrogen gas stream at 100 K. Form 2 crystals diffracted to 1.98 Å and belong to the space group $P6_522$, with unit cell parameters of $a = b$ = 91.5 Å, and c = 187.0 Å. Assuming the presence of one Csd3 monomer in the asymmetric unit, the Matthew's coefficient and the solvent content are 2.65 $Å^3$ Da^{-1} and 53.6 %, respectively.

The selenomethionine-substituted $Csd3_{\Delta 41}$ protein was not expressed in *E. coli*. Therefore, I prepared a Pt derivative of Form 2 crystals by soaking

them for 3 min in 2 μl of a heavy atom-containing cryoprotectant solution, which was the reservoir solution supplemented with 25% (*v/v*) glycerol and 30 mM K_2PtCl_4. Single-wavelength anomalous diffraction (SAD) data were collected at 100 K from the Pt-derivative crystal at a wavelength of 1.0720 Å using a ADSC Q270 CCD detector at the BL-7A experimental station of Pohang Light Source, Korea (Table 1-1). Raw data were processed and scaled using the program suit HKL2000 (Otwinowski and Minor, 1997). SAD phases were calculated using the program AUTOSOL of the PHENIX software package (Adams et al., 2010) and were further improved by RESOLVE (Terwilliger, 2003), yielding an interpretable electron density map at 2.95 Å. Phasing statistics are presented in Table 1-1.

Table 1-1. Statistics for data collection and phasing: *H. pylori* Csd3$_{\Delta 41}$

A. Data collection			
Data set	Form 1	Form 2	Pt (peak)
Space group	$P2_12_12_1$	$P6_522$	$P6_522$
Unit cell lengths, *a, b, c* (Å)	62.6, 112.1, 112.9	91.5, 91.5, 187.0	92.0, 92.0, 186.5
Unit cell angles, α, β, γ (°)	90, 90, 90	90, 90, 120	90, 90, 120
X-ray wavelength (Å)	1.0000	0.9000	1.0720
Resolution range (Å)	50.0–2.00 (2.03–2.00)[a]	50.0–1.98 (2.01–1.98)[a]	30.0–2.95 (3.00–2.95)[a]
Total / unique reflections	264,154 / 54,371	350,996 / 33,042	480,902 / 18,708[b]
Completeness (%)	99.5 (100.0)[a]	99.9 (100.0)[a]	100.0 (100.0)[a,b]
$\langle I \rangle / \langle \sigma_I \rangle$	27.5 (2.9)[a]	41.9 (3.5)[a]	64.0 (13.0)[a,b]
R_{merge}[c] (%)	8.3 (66.9)[a]	8.8 (87.7)[a]	12.9 (58.2)[a,b]
$CC_{1/2}$[d] (%)	99.7 (75.9)[a]	99.9 (91.6)[a]	99.9 (92.6)[a]
B. SAD phasing	Figure of merit (before / after density modification)		0.39 / 0.73

Footnotes to Table 1-1

[a] Values in parentheses refer to the highest resolution shell.

[b] Friedel pairs were treated as separate observations.

[c] $R_{merge} = \Sigma_h \Sigma_i | I(h)_i - \langle I(h) \rangle | / \Sigma_h \Sigma_i I(h)_i$, where $I(h)$ is the intensity of reflection h, Σ_h is the sum over all reflections, and Σ_i is the sum over i measurements of reflection h.

[d] $CC_{1/2}$ is the correlation coefficient of the mean intensities between two random half-sets of data.

1.2.4. Model building and refinement

The program RESOLVE was also used for auto-building, which resulted in an initial model accounting for ~38% of the residues in the recombinant polypeptide chain with much of the sequence assigned. The refined model of the Form 2 crystal was used as a search model to determine the $Csd3_{\Delta 41}$ structure in the Form 1 crystal by molecular replacement utilizing the program MOLREP (Vagin and Teplyakov, 2010). Manual model building was done using the program COOT (Emsley et al., 2010) and the models were refined with the programs REFMAC5 (Murshudov et al., 1997), including the bulk solvent correction. A total of 5% of the data was randomly set aside as test data for the calculation of R_{free} (Brünger, 1992). The stereochemistry of the refined models was assessed by MolProbity (Chen et al., 2010). Refinement statistics are presented in Table 1-2. The coordinates and structure factors have been deposited in the Protein Data Bank under accession codes 4RNY and 4RNZ for Form 1 and Form 2, respectively.

Table 1-2. Refinement statistics: *H. pylori* Csd3$_{\Delta 41}$

Data set	Form 1	Form 2
Resolution range (Å)	50.0–2.00	50.0–1.98
R_{work} / R_{free}[a] (%)	20.3 / 25.6	20.8 / 23.9
No. of non-hydrogen atoms / average *B*-factor (Å^2)		
Protein	5,822 / 37.0	2,919 / 40.4
Metal ion[b]	2 / 26.1	3 / 33.7
Water oxygen	225 / 34.6	148 / 40.8
Glycerol	30 / 49.3	18 / 55.0
Sulfate ion	45 / 42.3	-
Phosphate ion		15 / 65.0
R.m.s. deviations from ideal geometry		
Bond lengths (Å) / Bond angles (°)	0.013 / 1.58	0.009 / 1.42
Ramachandran[c]		
Favored / Outliers (%)	96.9 / 0.0	96.3 / 0.0
Poor rotamers (%)	0.16	0.00

Footnotes to Table 1-2

[a] $R_{work} = \Sigma \mid |F_{obs}| - |F_{calc}| \mid / \Sigma |F_{obs}|$, where R_{free} is calculated for a randomly chosen 5% of reflections, which were not used for structure refinement and R_{work} is calculated for the remaining reflections.

[b] Two metal ions in Form 1 are Zn^{2+} ions in the active site of two chains of Csd3$_{\Delta 41}$. Form 2 contains a Zn^{2+} ion in the active site and likely two Ni^{2+} ions bound to the C-terminal hexa-histidine tag.

[c] Values obtained using *MolProbity*.

1.2.5. Identification of Zn^{2+} binding by anomalous diffraction data

To test whether the metal-binding site is occupied by a Zn^{2+} ion, SAD data were collected from both Form 1 and Form 2 crystals at 100 K at the X-ray wavelength of 1.2820 Å at the beamline 5C of Pohang Light Source (Table 1-3). Anomalous difference maps were calculated using the FFT program of the CCP4i software package (Read and Schierbeek, 1988).

Table 1-3. Statistics for data collection of Zn SAD data sets

Data collection		
Data set	Zn SAD (Form 1)	Zn SAD (Form 2)
Space group	$P2_12_12_1$	$P6_522$
Unit cell lengths, a, b, c (Å)	62.2, 113.0, 112.7	92.4, 92.4, 186.7
Unit cell angles, α, β, γ (°)	90, 90, 90	90, 90, 120
X-ray wavelength (Å)	1.2820	1.2820
Resolution range (Å)	50.0–2.30 (2.34–2.30)[a]	50.0–2.35 (2.39–2.35)[a]
Total / unique reflections	559,489 / 68,296[b]	544,542 / 36,489[b]
Completeness (%)	100.0 (99.6)[a,b]	99.1 (97.4)[a,b]
$<I> / <\sigma_I>$	57.2 (3.6)[a,b]	65.4 (8.7)[a,b]
R_{merge}[c] (%)	8.8 (83.9)[a,b]	14.4 (72.4)[a,b]
$CC_{1/2}$[d] (%)	99.8 (85.3)[a]	99.5 (86.5)[a,b]

Footnotes to Table 1-3

[a] Values in parentheses refer to the highest resolution shell.

[b] Friedel pairs were treated as separate observations.

[c] $R_{merge} = \Sigma_h \Sigma_i | I(h)_i - < I(h) > | / \Sigma_h \Sigma_i I(h)_i$, where $I(h)$ is the intensity of reflection h, Σ_h is the sum over all reflections, and Σ_i is the sum over i measurements of reflection h.

[d] $CC_{1/2}$ is the correlation coefficient of the mean intensities between two random half-set of data.

1.2.6 Equilibrium sedimentation

Equilibrium sedimentation experiments were performed using a Beckman ProteomeLab XL-A analytical ultracentrifuge in buffer C at 4°C. The recombinant $Csd3_{\Delta 41}$ protein samples were monitored by measuring the absorbance at 280 nm using six-sector cells at three rotor speeds (12,000, 14,000, and 18,000 rpms) and three different protein concentrations (3.77, 5.03, and 6.29 μM). The protein concentration of the recombinant $Csd3_{\Delta 41}$ was estimated using ε_{280nm} = 39,770 M^{-1} cm^{-1} and a molecular mass of 42,562 Da, which includes the C-terminal hexahistidine-containing tag. Partial specific volume of the protein and the buffer density were calculated using Sednterp (Laue et al., 1992). Further data manipulation and data analysis by mathematical modeling were performed using MLAB (Knott, 1979).

1.3. Results and discussion

1.3.1. Structure determination of Csd3

H. pylori Csd3 was predicted to have a single transmembrane helix between Lys7 and Leu26, when its sequence was analyzed by the TMHMM Server v. 2.0 (Krogh et al., 2001). Therefore, I initially overexpressed the construct covering residues 29–403 fused with the C-terminal hexahistidine-containing tag in *E. coli* and the expressed protein was crystallized. However, the crystals diffracted poorly to low resolution only (~4 Å), despite extensive screening of crystallization conditions. This prompted me to try overexpressing a number of shorter constructs encompassing residues 40–403, 40–397, 40–393, 40–390, 42–403, 42–397, 42–393, 42–390, 45–403, 45–397, 45–393, and 45–390. Among them, only the N-terminal 41-residue truncated Csd3 ($Csd3_{\Delta 41}$) comprised of residues 42–403 was expressed in soluble form and it yielded two different forms (orthorhombic Form 1 and hexagonal Form 2) of well-diffracting crystals (Table 1-1).

I have determined the structure of $Csd3_{\Delta 41}$ (Figure 1-6) using the SAD data from a Pt-derivatized Form 2 crystal. The model of the Form 2 crystal was refined at 1.98 Å to R_{work} and R_{free} of 20.8 and 23.9%, respectively

(Table 1-2). The Form 2 crystal contains one monomer of $Csd3_{\Delta 41}$ in the asymmetric unit. This model of $Csd3_{\Delta 41}$ excludes six residues (Pro251–Arg256) on a disordered loop near the metal binding site as well as four residues Gly333–Thr336 and two histidines at the end of the C-terminal affinity tag. The model of the Form 1 crystal was refined at 2.00 Å to R_{work} and R_{free} of 20.3 and 25.6%, respectively (Table 1-2). The Form 1 crystal contains two monomers (chains A and B) of $Csd3_{\Delta 41}$ in the asymmetric unit. In both chains A and B, four residues Gly333–Thr336 and the C-terminal affinity tag (LEHHHHHH) are disordered. The two chains A and B are highly similar to each other, with an r.m.s. deviation of 0.67 Å for 359 C_α atoms. However, they show larger structural deviations from the chain of the Form 2 crystal, with r.m.s. deviations of 1.61 Å and 1.85 Å for 353 C_α atoms in each of chains A and B, respectively. Largest deviations occur in η1 and α6 helices, with 6.65 Å and 5.58 Å at C_α of Pro167 and Gly366, respectively (Figure 1-7). The observed structural variation is most likely due to the inherent flexibility of these regions and different crystal contacts.

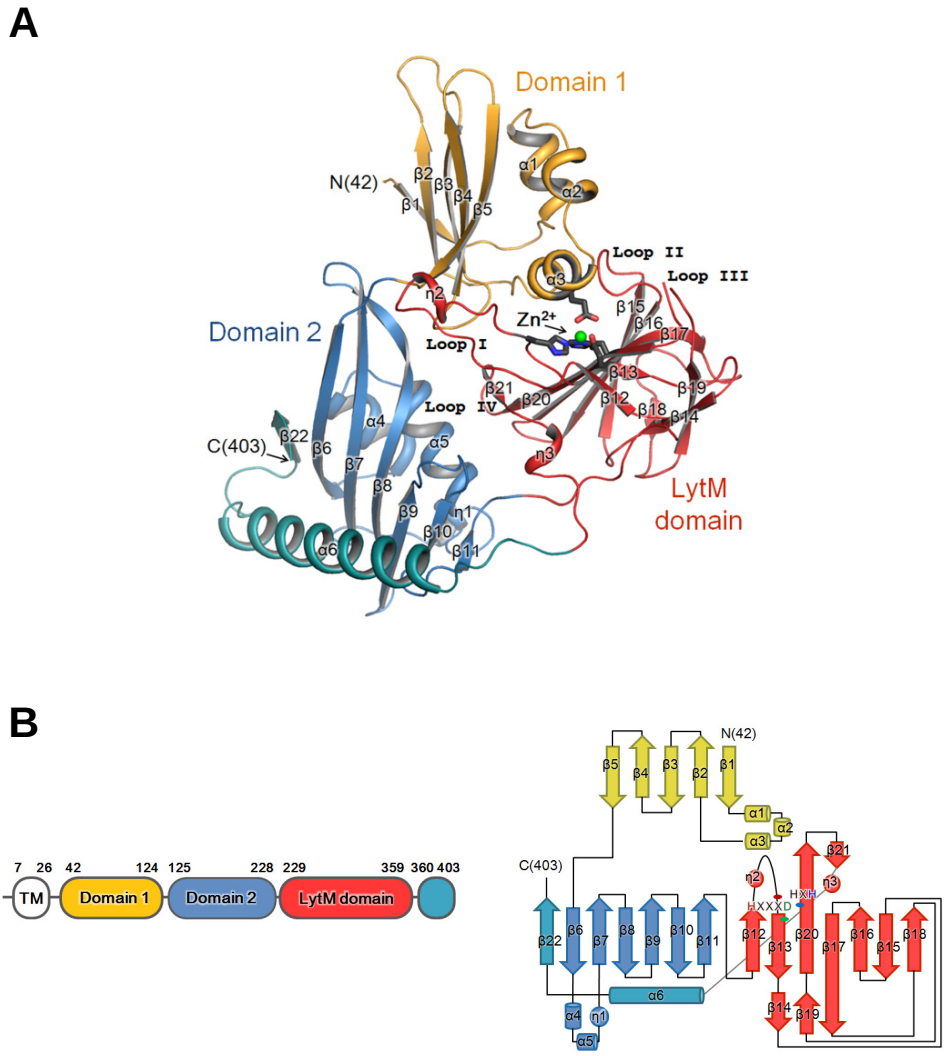


Figure 1-6. Overall monomer structure and topology of *H. pylori* Csd3$_{\Delta 41}$

(A) Ribbon diagram of the $Csd3_{\Delta 41}$ monomer, with the secondary structure elements labeled. Domain 1, the core of domain 2, and the LytM domain are shown in bright orange, skyblue, and red, respectively. The C-terminal α-helix (α6) and β-strand (β22) are colored in teal blue. A green sphere is a Zn^{2+} ion. Side chains of the metal-coordinating residues (Glu74, His259, Asp263, and His341) are shown in stick models (dark gray). The secondary structures have been defined by the STRIDE program (Heinig and Frishman, 2004). The walls of the active site in the LytM domain are made up of four loops: Loop I (β12-β13 loop), Loop II (β15-β16 loop), Loop III (β19-β20 loop), and Loop IV (β20-β21 loop). (B) Domains of *H. pylori* Csd3 (left). TM: transmembrane helix. Residue numbers for each domain are indicated. Topology diagram of $Csd3_{\Delta 41}$ (right). α-Helices, β-strands, and loops are shown as cylinders, arrows, and solid lines, respectively. Three 3_{10} helices are omitted for clarity. They are η1 between α5 and β7, η2 between β12 and β13, and η3 between β21 and α6. Structure figures were drawn using PyMOL (DeLano WL, 2002, The PyMOL Molecular Graphics System. DeLano Scientific, Palo Alto, California).

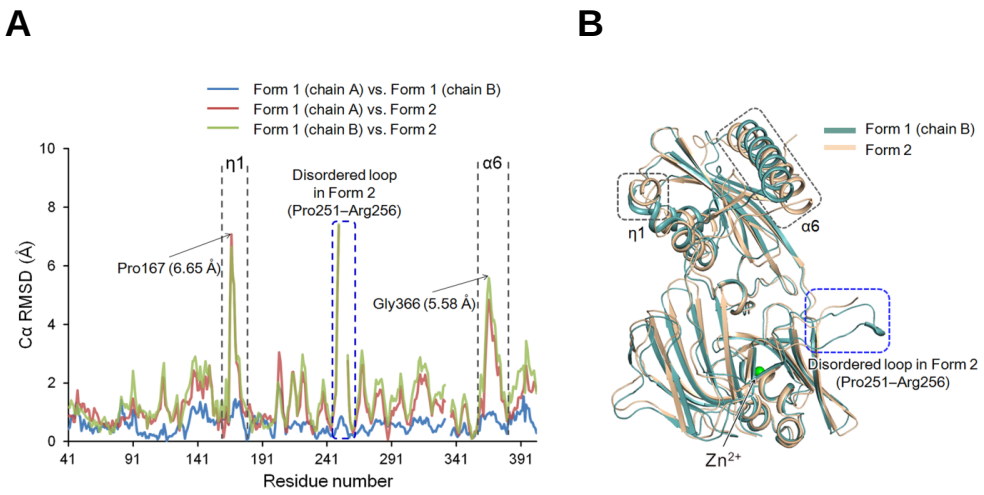


Figure 1-7. Comparisons of Csd3$_{\Delta 41}$ monomer structures

(A) Plot of Cα r.m.s. deviations for pairwise comparisons among three monomer models of Csd3$_{\Delta 41}$. Chains A and B of Form 1 crystal are more similar to each other than they are to Form 2 crystal. The η1 and α6 helices show large deviations between Form 1 and Form 2 crystals. (B) A superposition of two monomer models of Csd3$_{\Delta 41}$. Models of Form 1 crystal (chain B) and Form 2 crystal are colored in teal blue and light orange, respectively. The η1 and α6 helices, which show significant conformational differences, are highlighted by gray dotted boxes. Residues Pro251–Arg256 are disordered in Form 2 crystal and are highlighted by a blue dotted box.

1.3.2. Oligomeric state of $Csd3_{\Delta 41}$ in solution

In the Form 1 crystal, two monomers in the asymmetric unit bury a relatively large surface area of 1,130 $Å^2$ per monomer (6.0% of the monomer surface area), whereas the largest buried surface area in the Form 2 crystal is 710 $Å^2$ per monomer (3.8% of the monomer surface area). The bulk of the crystal interfaces have areas below 1,000 $Å^2$ with very few representatives above that value (Duarte et al., 2012). It is also well known that biological interfaces tend to exhibit large areas, with a majority of cases from 1,000 $Å^2$ and above (Duarte et al., 2012). This raised a question about the oligomeric state of $Csd3_{\Delta 41}$ in solution. Therefore, I carried out equilibrium sedimentation experiments at different rotor speeds and at different protein concentrations. All the measured data fit well to a homogeneous monomer model, indicating that $Csd3_{\Delta 41}$ exists as monomers in solution. A representative result measured at 18,000 rpm using a protein concentration of 3.77 μM is presented in Figure 1-8.

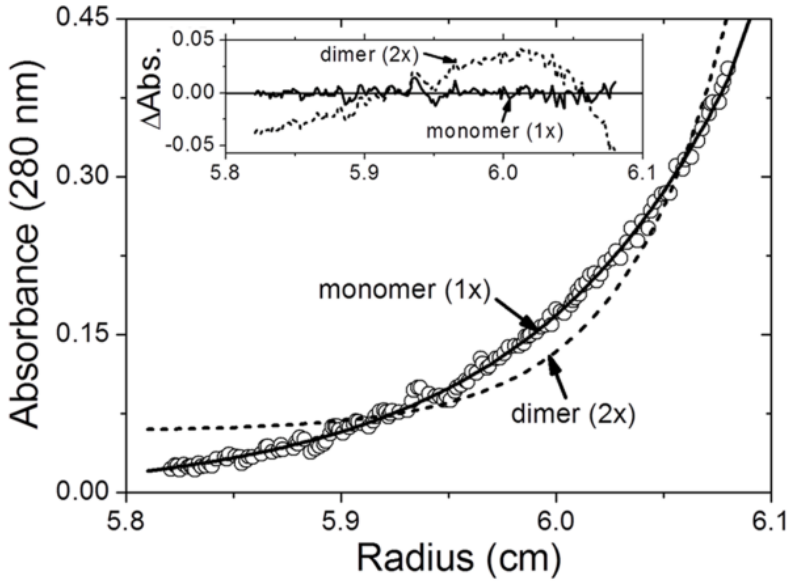


Figure 1-8. Equilibrium sedimentation distributions of *H. pylori* Csd3$_{\Delta 41}$

Representative data measured at 18,000 rpm using the 3.77 μM protein concentration are shown. The circles are experimental absorbance data at 280 nm and the solid line is a fitting line for a homogeneous monomer (1×) model. The dotted line is a fitting line for an ideal homogeneous dimer (2×) model. (Inset) Distributions of the residuals for monomer (1×, solid line) and dimer (2×, dotted line) models, respectively. Random distribution of the residuals for the monomer (1×) model indicates that Csd3$_{\Delta 41}$ exists as homogeneous monomers in solution.

1.3.3. Three domain structure of $Csd3_{\Delta 41}$ and structural similarity searches

The structure of $Csd3_{\Delta 41}$ can be divided into three domains: domain 1 (residues Glu42–Ile124), domain 2 (residues Ile125–Gly228 and Ala360–Phe403), and the LytM domain (residues Phe229–Thr359) (Figure 1-6). A structural similarity search using the Dali Server (Holm and Rosenström, 2010) revealed that the entire structure (Form 1 crystal, chain A) of $Csd3_{\Delta 41}$ resembles an outer-membrane protein from *Neisseria meningitidis* (NMB0315) (PDB code 3SLU; Wang et al., 2011) (an r.m.s. deviation of 5.2 Å for 315 equivalent Cα positions, a Z-score of 21.7, and a sequence identity of 26%) and a putative lysostaphin peptidase from *Vibrio cholera* (VC0503) (PDB code 2GU1; Ragumani et al., 2008) (an r.m.s. deviation of 4.5 Å for 219 equivalent Cα positions, a Z-score of 18.8, and a sequence identity of 23%) (Figure 1-9). The above two proteins are three-domain proteins, but domain 2 of VC0503 is not included in the structural overlap due to a large difference in domain arrangements. They belong to the M23B metallopeptidase family (Wang et al., 2011; Ragumani et al., 2008).

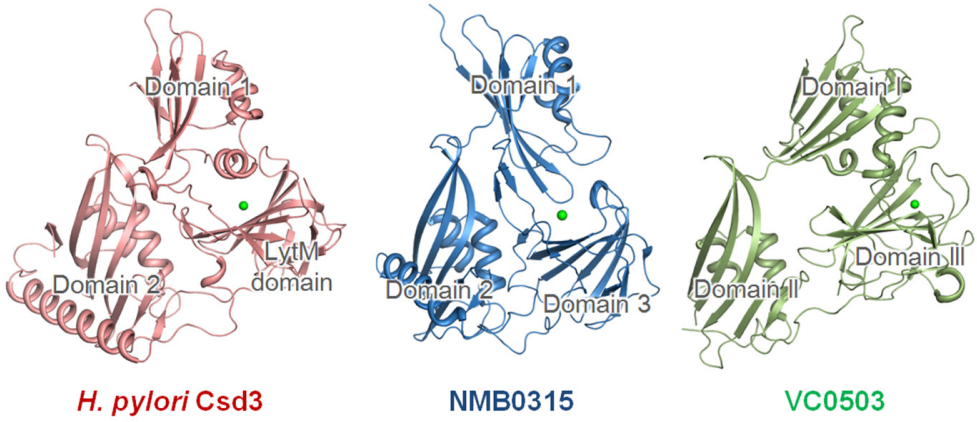


Figure 1-9. Three-domain proteins showing structural similarity to the entire structure of *H. pylori* Csd3$_{\Delta 41}$

The structure of *H. pylori* Csd3$_{\Delta 41}$ (deep salmon) is on the left, *N. meningitides* NMB0315 (sky blue; PDB code, 3SLU) in the middle, and *V. cholera* VC0503 (pale green; PDB code, 2GU1) on the right. Metal ions are shown as green spheres. In NMB0315, the Zn^{2+} ion was replaced with a Ni^{2+} ion during affinity chromatography (Wang et al., 2011).

Domain 1 of Csd3 has an α/β fold consisting of a β-sheet and three short α-helices (Figure 1-6B). The β-sheet consists of five anti-parallel β-strands that are arranged in the order β1↓-β2↑-β3↓-β4↑-β5↓. A cluster of continuously linked helices (α1-α2-α3) is inserted between strands β1 and β2, and is packed on one side of the β-sheet. The core of domain 2 (Ile125–Gly228) also has an α/β fold consisting of a β-sheet and three short helices (Figure 1-6B). The β-sheet consists of six anti-parallel β-strands arranged in the order β6↓-β7↑-β8↓-β9↑-β10↓-β11↑, with a long, highly twisted β8 strand in the middle. As in domain 1, a cluster of consecutively-linked helices (α4-α5-η1) is inserted between strands β6 and β7, and lies on the concave side of the β-sheet. Domain 1 and the core of domain 2 share a common fold despite a very low level of amino-acid sequence identity (only 15 residues are identical). However, their charge distributions on the surface are distinct from each other (Figure 1-10). When the Dali search was performed separately against domain 1 (Form 1 crystal, chain A), the highest structural similarity is observed with the corresponding domains from VC0503 and NMB0315, with Z-scores of 7.6 and 5.0, respectively. Besides these, domain 1 displays a very remote structural similarity with single-chain monellin (PDB code 1MNL; Lee et al., 1999) with a Z-score of 4.2 and

human cystatin A (also called stefin A) (PDB code 3KSE, chain D; unpublished deposition) with a Z-score of 3.5. When the Dali search was performed separately against the core of domain 2 (Form 1 crystal, chain A), the highest structural similarity is observed with the corresponding domains from VC0503 and NMB0315, with Z-scores of 14.6 and 14.5, respectively. Besides these, domain 2 exhibits a very remote structural similarity with the core of a single-stranded DNA-binding protein from *Aeropyrum pernix* K1 (PDB code 4PSO; Ghalei et al., 2014) and perfringolysin O from *Clostridium perfringens* (PDB code 1M3I; unpublished deposition) with Z-scores of 4.5 and 4.4, respectively.

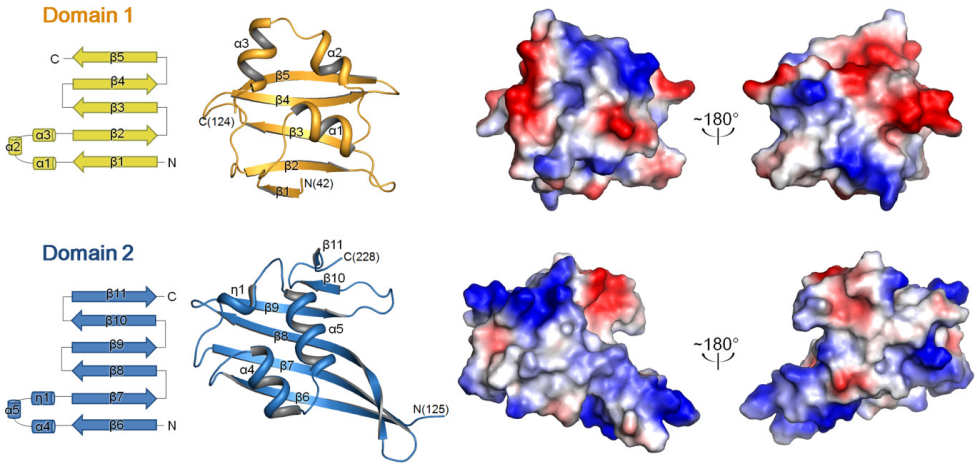


Figure 1-10. Comparison of charge distributions on the surface of Csd3 domain 1 (top) and domain 2 (bottom)

Topology and ribbon diagrams for domain 1 and domain 2 are colored as in Figure 1-6A, with the secondary structure elements labeled. Two different views of the electrostatic potential surface diagrams related by a 180° rotation are shown next to the ribbon diagrams.

The bulk of Csd3 LytM domain is folded as a two-layered sandwich of a larger, seven-stranded anti-parallel β-sheet (β12↑-β13↓-β20↑-β17↓-β16↑-β15↓-β18↑) and a smaller, three-stranded anti-parallel β-sheet (β14↓-β19↑-β17↓). These β-sheets share a long strand (β17), which is bent like the letter J. Another long strand (β20) associates with a short strand (β21) to form a mini, anti-parallel β-sheet (β20↑-β21↓), which is followed by a 3_{10} helix (η3) (Figure 1-6B). As expected, the LytM domain (Form 1 crystal, chain A) of $Csd3_{\Delta 41}$ exhibits significant structural similarities with the corresponding domains of the M23 metallopeptidases, such as NMB0315 (PDB code 3SLU; an r.m.s. deviation of 2.8 Å for 123 equivalent Cα positions, a Z-score of 19.7, and a sequence identity of 44%) and VC0503 (PDB code 2GU1; an r.m.s. deviation of 2.7 Å for 120 equivalent Cα positions, a Z-score of 19.0, and a sequence identity of 35%). It is also structurally similar to the LytM domains of *Staphylococcus aureus* glycylglycine endopeptidase, another M23B metallopeptidase (PDB code 2B13; Firczuk et al., 2005) (an r.m.s. deviation of 1.9 Å for 118 equivalent Cα positions, a Z-score of 18.3, and a sequence identity of 30%) and virulence factor LasA from *Pseudomonas aeruginosa*, an M23A metallopeptidase (PDB code 3IT7; Spencer et al., 2010) (an r.m.s. deviation of 2.1 Å for 107 equivalent Cα

positions, a Z-score of 11.4, and a sequence identity of 21%). M23A metallopeptidases are distinguished from the more numerous M23B enzymes, by structural features that include disulfide bridges and possession of an additional C-terminal subdomain as well as alterations to the active site region that are manifest in differences in the sequence spacing between His and Asp residues of the HxxxD motif of M23B family (Spencer et al., 2010). In M23A metallopeptidases, the intervening sequence is variable and several times longer than that of M23B.

After leaving the LytM domain, the polypeptide chain is connected to a C-terminal α-helix (α6) and a short β-strand (β22) that fold back onto the core of domain 2, making a tight interaction with the core of domain 2 (Figure 1-11). The α6 helix sits on the opposite side of the β-sheet of domain 2 from the cluster of three helices (α4-α5-η1) and is tightly anchored to the β-sheet through hydrophobic and hydrogen bond interactions (Figure 1-11). The β22 strand lies antiparallel to the β6 strand of domain 2 and forms the seventh strand of the β-sheet in the core of domain 2 (Figure 1-11). The C-terminal α6–β22 region may play an important role in stabilizing the inter-domain orientation of the core of domain 2 and the LytM domain of Csd3.

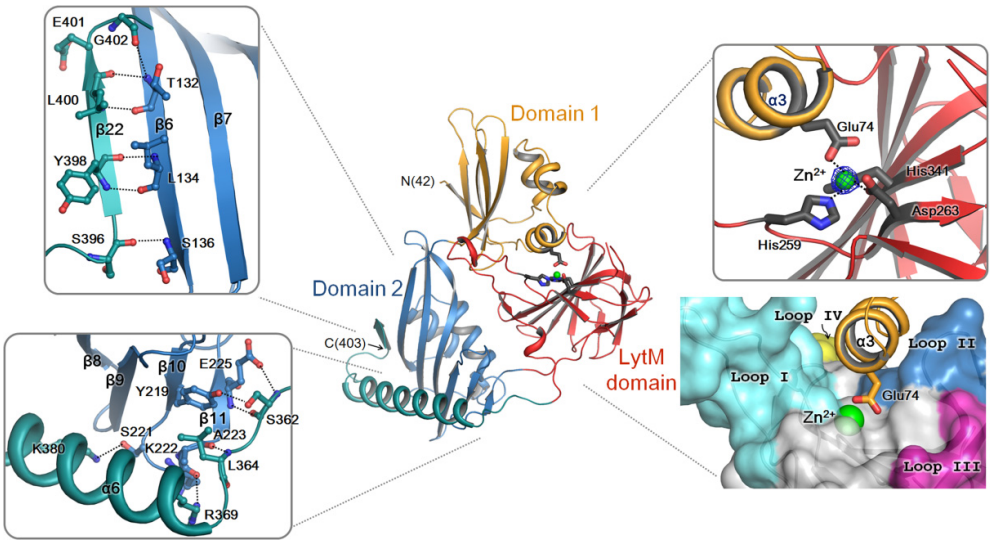


Figure 1-11. Metal coordination in $Csd3_{\Delta 41}$ and interactions of the C-terminal α-helix (α6) and β-strand (β22) with the core of domain 2

Ribbon diagram of the $Csd3_{\Delta 41}$ monomer, colored as in Figure 1-6A, is shown in the center. Close-up views on the left represent interactions of the C-terminal strand (β22) with the β6 strand in the core of domain 2 (top) and the C-terminal helix (α6) with the β-sheet in the core of domain 2 (bottom). The hydrogen bond interactions are shown in black dotted lines. Close-up views on the right represent the ribbon diagram of the Zn^{2+}-binding motif (top) and the surface representation of the substrate-binding groove formed by four loops of the LytM domain (bottom). The electron density for the Zn^{2+} ion in the omit $mF_o - DF_c$ map is shown in blue colored mesh (contoured at 10σ). To show the detailed interactions more clearly, close-up views have slightly different orientations.

1.3.4. Active site of the Csd3 LytM domain

The Csd3 LytM domain possesses characteristic features of the MEROPS M23 family of metallopeptidases. In these enzymes, the catalytic residues are anchored by the central antiparallel β-sheet and are grouped around a metal ion. I confirmed that this metal-binding site is indeed occupied by a Zn^{2+} ion by calculating anomalous difference maps using the anomalous data collected at the zinc absorption edge of 1.2820 Å (Table 1-3 and Figure 1-12). In all $Csd3_{\Delta 41}$ structures, I observed a tetra-coordinated Zn^{2+} ion in the active site, with the expected three amino-acid ligands (His259$^{\varepsilon 2}$, Asp263$^{\delta 1}$, and His341$^{\delta 1}$) and one amino-acid ligand (Glu74$^{\varepsilon 2}$) from the α3 helix of domain 1 (Figures 1-11 and 1-12). His259 and Asp263 belong to the characteristic **H**xxx**D** motif, while His341 is the second histidine of the Hx**H** motif (Figure 1-13). The metal-ligand atom distances are in the range of 1.95–2.11 Å, consistent with typical Zn^{2+} ion-ligand atom distances. If Csd3 were in the active state for the peptidase activity, two water molecules should occupy the positions close to the side-chain oxygen atoms of Glu74, as penta-coordination is believed to be consistent with proposals for the catalytic mechanism (Sabala et al., 2014). The distances between Zn^{2+} and two oxygens from the side chain of Glu74 (Glu74$^{\varepsilon 1}$ and

Glu74$^{\varepsilon 2}$) are 2.9 Å and 2.0 Å, respectively. This coordination varies from that in the uncomplexed structure of LasA from *P. aeruginosa* (Spencer et al., 2010). The Zn^{2+} ion in the uncomplexed structure of LasA is described as penta-coordinated (in a slightly distorted trigonal bipyramidal geometry) with conserved three metal ligands and two water molecules. The distances between Zn^{2+} and the water oxygen atoms (Wat-1 and Wat-2) are 2.1 Å and 2.7 Å, respectively. In the tartrate-complexed structure of LasA, tartrate oxygen atoms occupy nearly identical positions of both Zn^{2+}-coordinated water molecules, with the distances between Zn^{2+} and tartrate oxygen atoms are 2.5 Å and 2.1 Å, respectively. However, Zn^{2+}-coordination in the tartrate-complexed structure was described as tetrahedral geometry (Spencer et al., 2010). It has been suggested that penta-coordination is consistent with many proposals for catalytic mechanism and that observations of variable Zn^{2+}-coordination among the characterized M23 family metallopeptidases may reflect the low energy barriers for changes in Zn^{2+} coordination for catalysis (Sabala et al., 2014; Spencer et al., 2010).

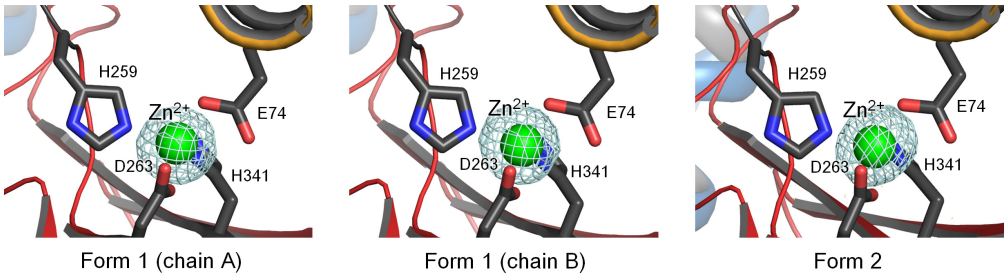


Figure 1-12. Identification of the Zn^{2+}-binding in the metal-binding site

Electron density of Zn^{2+} ions bound to the metal-binding site in the anomalous difference Fourier maps (contoured at 10σ and colored in cyan), calculated using the Zn SAD data sets from Form 1 and Form 2 crystals (Table 1-3). Ribbon diagrams are colored as in Figure 1-6A.

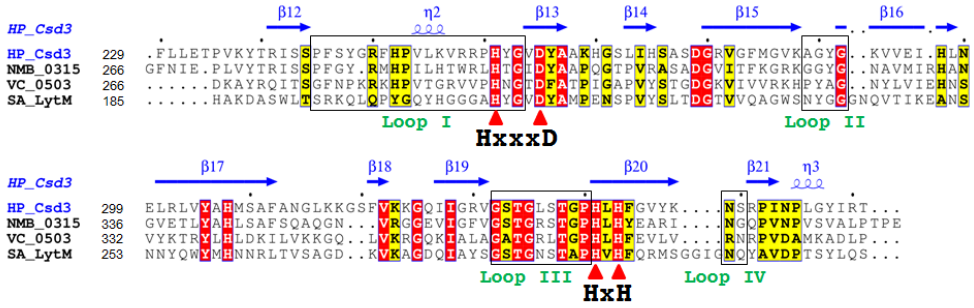


Figure 1-13. Sequence alignment of LytM domains of four M23B metallopeptidase proteins

Sequence alignment of Csd3 from *Helicobacter pylori* strain 26695 (HP0506; HP_Csd3; SWISS-PROT accession code, O25247), NMB0315 from *Neisseria meningitidis* (NMB_0315; Q9K163), VC0503 from *Vibrio cholerae* (VC_0503; Q9KUL5), and LytM from *Staphylococcus aureus* (SA_LytM; O33599). Red triangles indicate the conserved residues in the HxxxD and HxH motifs (H259-xxx-D263 and H339-x-H341 in *H. pylori* Csd3) that are important for the metallopeptidase activity. Four loops are indicated by black boxes.

The floor of the substrate-binding groove of the LytM domain is built by the central β-sheet and the walls of the active site are made up of four loops: Loop I (β12-β13 loop), Loop II (β15-β16 loop), Loop III (β19-β20 loop), and Loop IV (β20-β21 loop) (Figures 1-11 and 1-14). A superimposition of the LytM domain of Csd3 with those of NMB0315, VC0503, and *S. aureus* LytM reveals that the central β-sheet has little deviation with each other, but the difference is larger for the Loops I–IV (Figure 1-14B). Loop I of Csd3 shows the most significant deviation among the homologs. In Form 1 crystals, it is involved in crystal packing interactions by forming sulfate-mediated salt bridges (Figure 1-15). In Form 2 crystals, it is not involved in the crystal packing and is disordered. Therefore, I conclude that the observed structural difference of Loop I is largely due to the difference in crystal packing and also due to its inherent flexibility. In all structures of $Csd3_{\Delta 41}$, three residues (Gly333-Leu334-Ser335) of Loop III are disordered. Loop III has significant sequence conservation and immediately precedes the HxH motif (Figure 1-13). Gly333 and Ser335 of Csd3 are conserved, whereas Leu334 in Csd3 is substituted by Arg368 and Arg365 in NMB0315 and VC0503, respectively, and Asn286 in *S. aureus* LytM (Figure 1-13). In the structure of the active LytM from *S. aureus* (PDB: 2B44), a phosphate ion, as a substrate candidate,

is fixed in space in the active site by multiple hydrogen bonds with Asn286 (Firczuk et al., 2005). Ser369 of NMB0315 and Ser287 of the active LytM from *S. aureus*, corresponding to the disordered Ser335 on Loop III of Csd3, interact with the first histidine of the HxH motif. This histidine is considered as a catalytic residue by coordinating a Zn^{2+}-bound water molecule that acts as the nucleophile in the hydrolytic reaction (Spencer et al., 2010). Interestingly, the side chain of His339 of Csd3, first histidine of the HxH motif, shows two different orientations. Its orientation in Form 2 crystals is similar to those in other homologous proteins (Figure 1-16). However, it is flipped about 104° toward the solvent in Form 1 crystals (Figure 1-16A).

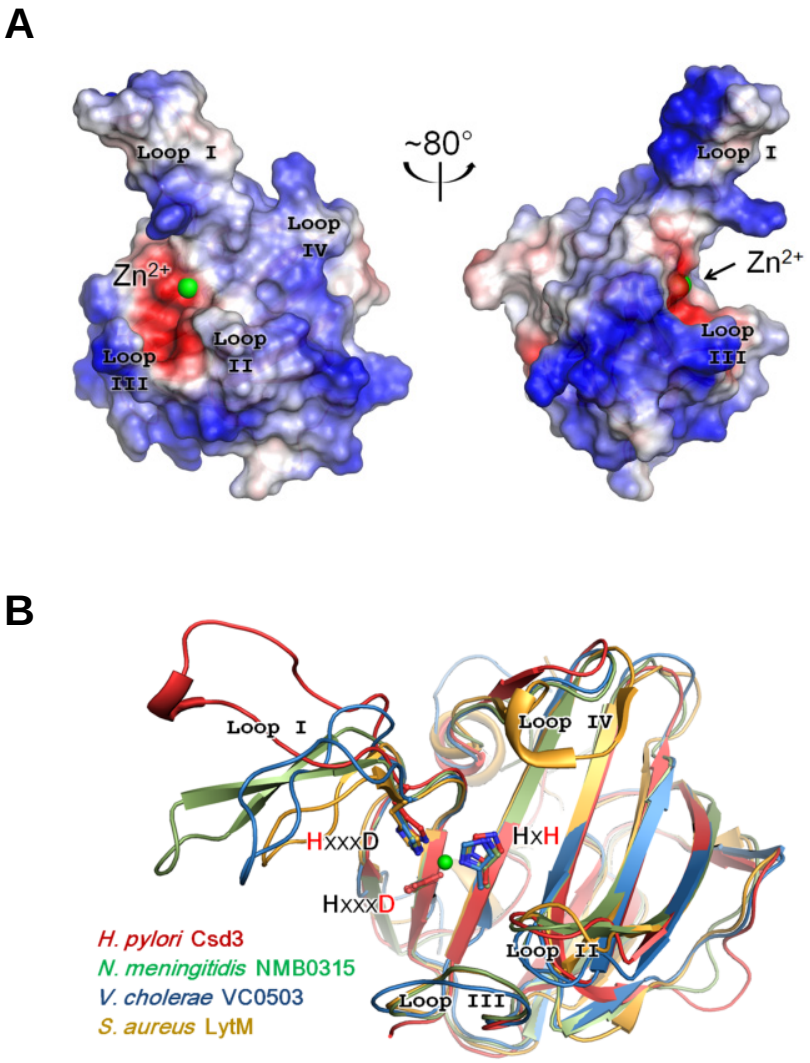


Figure 1-14. LytM domain of Csd3

(A) Two different views of the electrostatic potential surface diagrams of the LytM domain of Csd3. The positive and negative electrostatic potentials on the surface are colored in blue and red, respectively. Four loops that form the substrate-binding groove around the Zn^{2+} ion (green sphere) are denoted by Loops I–IV. (B) Superposition of LytM domains in four M23B metallopeptidases. LytM domains of *H. pylori* Csd3 (red), *N. meningitides* NMB0315 (pale green; PDB code, 3SLU), *V. cholera* VC0503 (sky blue; PDB code, 2GU1), and *S. aureus* LytM (yellow-orange; PDB code, 1QWY) are superimposed and shown in ribbon diagrams.

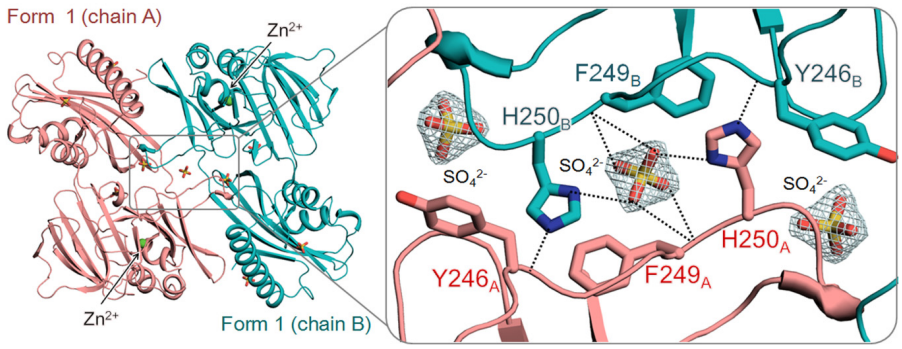


Figure 1-15. Crystal packing interactions in Form 1 crystal of *H. pylori* Csd3$_{\Delta 41}$

Chain A (salmon) and chain B (teal blue) of Csd3$_{\Delta 41}$ in the asymmetric unit are shown in ribbon diagrams. Zn^{2+} ions are indicated by green spheres. Nine sulfate ions are shown in stick models. A close-up view on the right represents inter-chain interactions mediated by a sulfate ion at the noncrystallographic two-fold symmetry axis. Residues at the interface between two chains are shown in stick models and labeled. The electron density for the sulfate ions in the $2mF_O - DF_C$ map are shown in cyan colored mesh (contoured at 2σ). Hydrogen bonds and salt bridges are indicated by black dotted lines.

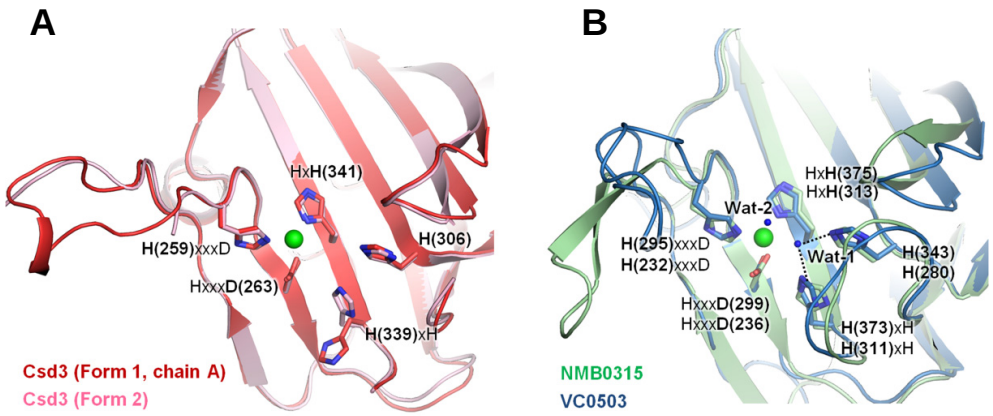


Figure 1-16. Conserved residues in the active site of M23B metallopeptidases

(A) Superposition of LytM domains in two crystal forms of *H. pylori* Csd3. Chain A of Form 1 crystal (red) and Form 2 crystal (pink) are shown in ribbon diagrams. (B) Superposition of LytM domains in two M23B family members (NMB0315 and VC0503), colored as in Figure 1-14B. Conserved residues are shown in stick models and are labeled (NMB0315 at the top and VC0503 at the bottom). Metal ions and water molecules are shown as green spheres and blue dots, respectively. In NMB0315, the Zn^{2+} ion was replaced with a Ni^{2+} ion during affinity chromatography (Wang et al., 2011). Wat-1 and Wat-2 are present in NMB0315, where the Ni^{2+} ion is penta-coordinated. Black dotted lines denote hydrogen bonds to Wat-1 in NMB0315. In VC0503, the Zn^{2+} ion is tetra-coordinated, with one water molecule (omitted for clarity) located between Wat-1 and Wat-2.

1.3.5. Domain 1 occludes the active site of LytM domain

In both Form 1 and Form 2 crystals, domain 1 blocks the active-site cleft of the LytM domain (Figure 1-17), with the protruding helix α3 contributing to the Zn^{2+} coordination sphere. Several negatively charged residues from domain 1 (Asp72, Glu74, and Glu78) and positively charged residues (Arg257, Arg301, and Arg349) as well as the Zn^{2+} ion from the LytM domain form strong salt bridges at the interface between these domains. As described above, the Zn^{2+} ion in the active site of the LytM domain is tetrahedrally coordinated by the three conserved residues (His259 and Asp263 of the HxxxD motif and the second histidine His341 of the HxH motif) of the LytM domain as well as non-conserved Glu74 on helix α3 of domain 1. Glu74 and two other acidic residues Asp72 and Glu78 of domain 1 are located at the domain interface and interact with the LytM domain through salt bridges and hydrogen bonds (Figure 1-17). There is a hydrogen bond between Glu74 and the conserved His306 of the LytM domain. Asp72 interacts with Arg257 on Loop I of the LytM domain by forming a salt bridge. Glu78 forms a salt bridge with Arg301 on the β17 strand of the LytM domain; it also makes a hydrogen bond with Tyr289 on Loop II of the LytM domain. In addition, Asp105 on the β3-β4 loop of domain 1 forms a salt

bridge with Arg349 on the β21 strand of the LytM domain. Tyr65 on α2 helix of domain 1 forms a hydrogen bond with the main-chain nitrogen atom of Gly288 on Loop II of the LytM domain. Ser103 on the β3-β4 loop of domain 1 is hydrogen bonded to Tyr260 on the β13 strand of the LytM domain.

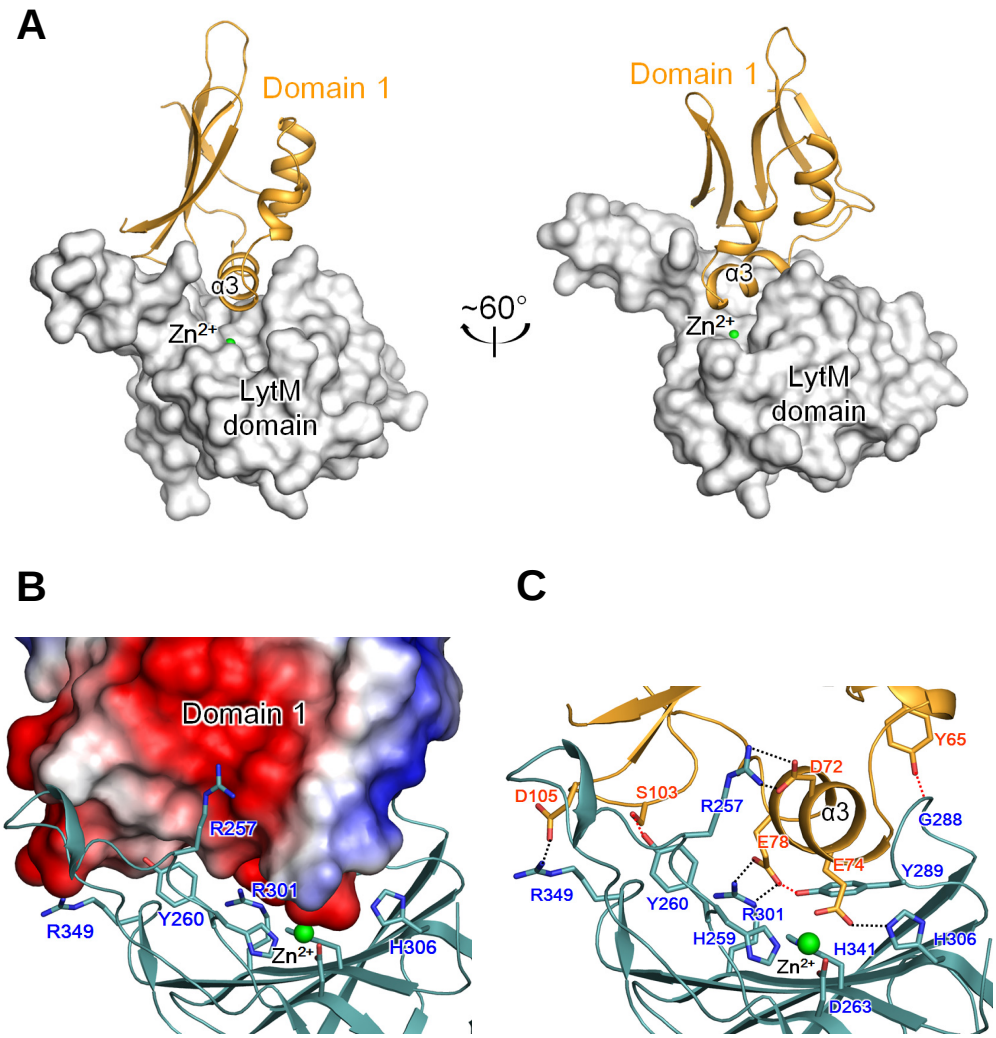


Figure 1-17. The active site in the Csd3 LytM domain is blocked by domain 1

In this figure, domain 2 is omitted for clarity. (A) Two different views of the interaction between domain 1 (shown in bright orange ribbons) and LytM domain (shown in a gray surface diagram) of Csd3. (B) Electrostatic surface diagram of domain 1 and ribbon diagram of LytM domain (colored in deep teal blue). Residues of LytM domain located at the domain interface are shown in stick models and are labeled. (C) Ribbon diagram of domain 1 (in bright orange) and LytM domain (colored in deep teal blue). Residues at the domain interface are shown in stick models. The hydrogen bond interactions and salt bridge interactions are shown in red and black dotted lines, respectively.

On the basis of these extensive interactions between domain 1 and the LytM domain, I suggest that $Csd3_{\Delta 41}$ in the crystal is in an auto-inhibited state. This is similar to the occlusion of the active site by an N-terminal segment in structurally related three-domain proteins in the inhibited conformational state (Figure 1-18). In NMB0315, the N-terminal short β3-β4 loop stretches into the active site and tightly associates with the catalytic domain (Wang et al., 2011). In VC0503, an N-terminal helix (α2) occupies the active-site cleft (Ragumani et al., 2008). In both cases, the loop or helix does not participate in metal coordination.

The Zn^{2+} ion bound to the active site of Csd3 is tetrahedrally coordinated by the side chains of four residues (Glu74, His259, Asp263, and His341) without any ordered water molecules. Glu74 is not conserved, whereas His259, Asp263, and His341 are conserved (Figure 1-13). The ligand position corresponding to the non-conserved Glu74 is usually occupied by a catalytic water molecule or various anions in the active forms of other M23 metallopeptidases (Figure 1-18). In the case of LytM from *S. aureus*, a truncated version that lacks Asn117 has much higher specific activity than the full-length enzyme in which poorly conserved Asn117 of the inhibitory domain occupies the one of the ligand sites of Zn^{2+} (Odintsov et al., 2004). In the active, truncated LytM from *S. aureus*, Asn117 is

replaced by phosphate, cacodylate, or tartrate from the crystallization buffer (Firczuk et al., 2005).

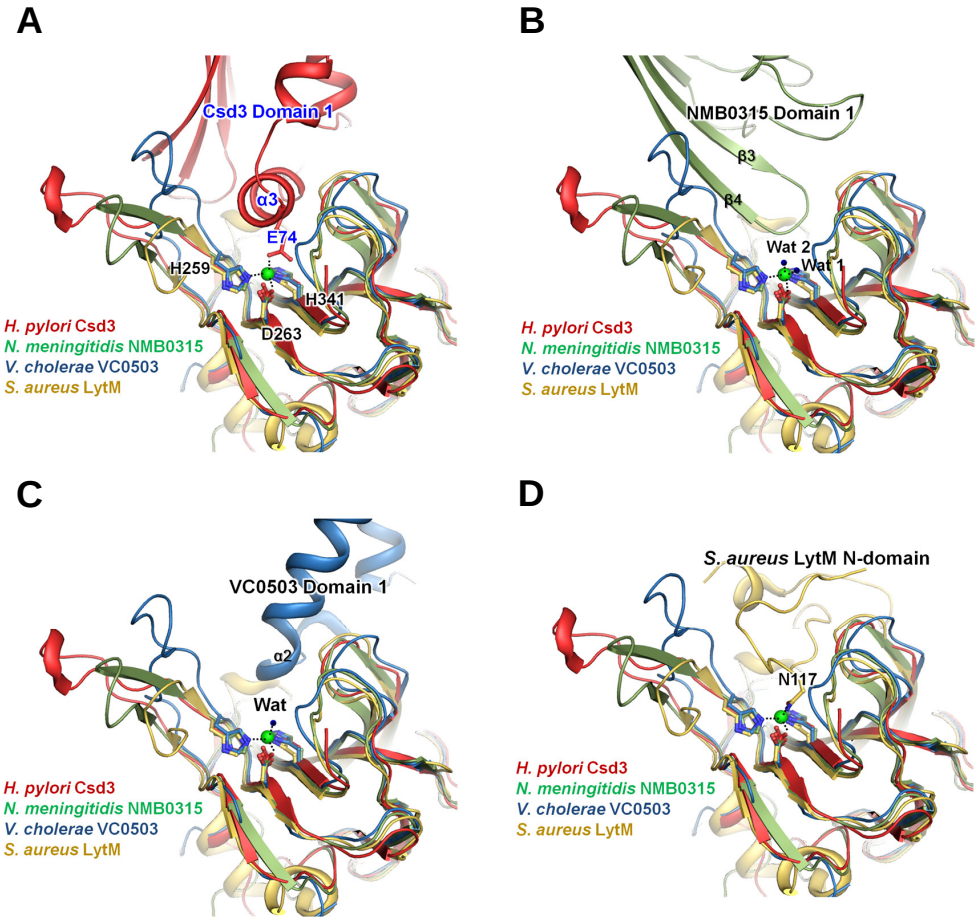


Figure 1-18. Superposition of LytM domains of M23B family metallopeptidases in the inhibited state

Four LytM domains are superimposed as in Figure 1-14B but are shown in a different orientation. (A–D) Domain 1 of *H. pylori* Csd3 (A), and corresponding inhibitory domains of *N. meningitidis* NMB0315 (B), *V. cholerae* VC0503 (C), and *S. aureus* LytM (D) are shown in addition to the superposed LytM domains. Metal ions are shown as green spheres. In NMB0315, the Zn^{2+} ion was replaced with a Ni^{2+} ion during affinity chromatography (Wang et al., 2011). Water molecules are shown as blue dots. Black dotted lines denote metal coordination.

1.3.6. Discussion

I have determined the first crystal structure of Csd3 from *H. pylori*, a protein that influences the helical cell shape crucial for the survival of *H. pylori* in the stomach. It is worth mentioning that the fold of Csd3 domain 1 is very remotely related to monellin/cystatin superfamily proteins, some of which act as inhibitors against the cysteine peptidases. The active site of the C-terminal LytM domain is blocked by the inhibitory domain 1, in particular, by helix α3 with the Zn^{2+}-coordinating Glu74. The structure of an AmiB ortholog from *Bartonella henselae* revealed that the active site of AmiB is similarly occluded by a conserved α-helix with a Zn^{2+}-coordinating glutamate residue (Glu290) and it was suggested that auto-inhibition is a critical feature of the regulation of peptidoglycan amidases required for cell division in Gram-negative bacteria (Yang et al., 2012). When I measured the CPase activity of *H. pylori* $Csd3_{\Delta 41}$ against the synthetic muramyl pentapeptide (the substrate for the CPase activity of Csd3) in the presence of Zn^{2+} using mass analysis, no reaction product was observed in all the reaction conditions that I tested. This result is consistent with the latent, inactive state of $Csd3_{\Delta 41}$ in the crystal.

This work suggests that the inhibitory domain 1, including the α3 helix, should be displaced from the active site of the LytM domain for the activation of the latent Csd3. The activation of *H. pylori* Csd3 may occur by autoproteolysis or may require proteolytic cleavage by other endopeptidases to free the catalytic LytM domain from the inhibitory domain 1. When I tried limited proteolysis with trypsin, chymotrypsin, and pepsin, no discrete cleavage at specific sites took place. To physically remove the inhibitory domain 1 from Csd3, I tried to overexpress various constructs (residues 80–403, 81–403, 83–403, 124–403, 125–403, 127–403, 130–403, 227–403, 230–403, 233–403, 125–359, 127–359, 227–359, 230–359, and 233–359) that lack domain 1. However, none of them was expressed in *E. coli* in a soluble form. Alternatively, a large conformational changed induced by an allosteric regulator may lead to the opening of the active site. For instance, a large conformational change culminates in opening of the active site to permit substrate entry, when an allosteric site ~60 Å distant from the D,D-transpeptidase active site is occupied in penicillin binding protein 2a (PBP2a) from *S. aureus* (Otero et al., 2013).

To provide details on the interaction of the substrate with Csd3, I have attempted soaking the Form 1 and Form 2 crystals in a cryoprotectant

solution that was supplemented with the muramyl pentapeptide. No electron density of the substrate peptide was observed. I have also tried to co-crystallize $Csd3_{\Delta 41}$ in the presence of excess muramyl pentapeptide under the conditions similar to the crystallization conditions of Form 1 and Form 2 crystals but this attempt did not produce any crystal. Due to a limited supply of the muramyl pentapeptide, it was not possible to perform more extensive screening of the crystallization conditions in its presence. Co-crystallization of $Csd3_{\Delta 41}$ in the presence of crosslinked dimers of muramyl penta- and tetra-peptides (the substrate for the EPase activity of Csd3) was not possible, because this substrate is neither commercially available nor easy to synthesize.

H. pylori Csd1 (HP1543), Csd2 (HP1544), and Csd3 (HP0506) were identified to contain the LytM domain of the M23B metallopeptidase family; they all may catalyze the same reaction as EPases. Csd2 could be a non-peptidase homolog of M23B family, as the first histidines of the characteristic HxxxD and HxH motifs in its expected LytM domain are mutated into Glu at residue 165 and Lys at residue 246, respectively. Csd3 exhibits the CPase activity, while Csd1 may or may not have such an activity. *H. pylori* Csd3 (with 403 residues) is considerably longer in sequence compared to Csd1 (312 residues) and Csd2 (308 residues). The significantly

shorter length of Csd1 and Csd2 indicates the presence of only one domain in their proregions. This may affect their CPase activity. However, no information on the three-dimensional structures of Csd1 and Csd2 is available at present and thus structural studies on them are required to compare the structures of Csd1 and Csd2 with Csd3 and to understand the functional differences in structural terms.

Bacteria may change their morphology to fit the circumstances for survival (Young, 2007). Under stressed conditions such as sub-inhibitory concentrations of antibiotics, *H. pylori* is able to enter the viable but non-culturable state, in which the microorganism modifies its morphology from a spiral to a coccoid form (Cellini, 2014). The viable coccoid form is more persistent in the host and environment (Cellini, 2014). It has been reported that overproduction of Csd3/HdpA in *H. pylori* N6 strain led to the transformation from rod-shaped to viable cocci-shaped bacteria (Bonis et al., 2010). This raises the possibility that transition of the helical to coccoid shape may be associated with the overexpression of Csd3. If this is the case, Csd3 could be an attractive drug target not only for eradicating the helical-shaped *H. pylori* but also for inhibiting the morphological transformation into the persistent coccoid form.

1.4. References

Ayala, G., Escobedo-Hinojosa, W. I., de la Cruz-Herrera, C. F., and Romero, I. (2014) Exploring alternative treatments for *Helicobacter pylori* infection. World J Gastroenterol. **20**:1450–1469.

Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. sect. D Biol. Crystallogr. **66:**213–221.

Bauer, B. and Meyer, T. F. (2011) The human gastric pathogen *Helicobacter pylori* and its association with gastric cancer and ulcer disease. Ulcers **2011**:23 pages. (DOI: 10.1155/2011/340157)

Berg, H. C., and Turner, L. (1979) Movement of microorganisms in viscous environments. Nature **278**:349–351.

Bonis, M., Ecobichon, C., Guadagnini, S., Prévost, M. C., and Boneca, I. G. (2010) A M23B family metallopeptidase of *Helicobacter pylori* required for cell shape, pole formation and virulence. Mol. Microbiol. **78:**809–819.

Brünger, A. T. (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature **355:**472–474.

Cellini, L. (2014) *Helicobacter pylori*: a chameleon-like approach to life. World. J. Gastroenterol. **20:**5575–5582.

Chan, A. C., Blair, K. M., Liu, Y., Frirdich, E., Gaynor, E. C., Tanner, M. E., Salama, N. R., and Murphy, M. E. (2015) Helical shape of *Helicobacter pylori* requires an atypical glutamine as a zinc ligand in the carboxypeptidase Csd4. J. Biol. Chem. **290**:3622–3638.

Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. sect. D Biol. Crystallogr. **66**:12–21.

Costa, K., Bacher, G., Allmaier, G., Dominguez-Bello, M. G., Engstrand, L., Falk, P., de Pedro, M. A., and García-del Portillo, F. (1999) The morphological transition of *Helicobacter pylori* cells from spiral to coccoid is preceded by a substantial modification of the cell wall. J. Bacteriol. **181:**3710–3715.

Duarte, J. M., Srebniak, A., Schärer, M. A., and Capitani, G. (2012) Protein interface classification by evolutionary analysis. BMC Bioinformatics **13**, 334.

Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. Acta Crystallogr. sect. D Biol. Crystallogr. **66:**486–501.

Firczuk, M., Mucha, A., and Bochtler, M. (2005) Crystal structures of active LytM. J. Mol. Biol. **354:**578–590.

Ghalei, H., von Moeller, H., Eppers, D., Sohmen, D., Wilson, D. N., Loll, B., and Wahl, M. C. (2014) Entrapment of DNA in an intersubunit tunnel system of a single-stranded DNA-binding protein. Nucleic Acids Res. **42:**6698–6708.

Hazell, S. L., Lee, A., Brady, L., and Hennessy, W. (1986) *Campylobacter pyloridis* and *gastritis*: association with intercellular spaces and adaptation to an environment of mucus as important factors in colonization of the gastric epithelium. J. Infect. Dis. **153**:658–663.

van Heijenoort, J. (2011) Peptidoglycan hydrolases of *Escherichia coli*. Microbiol. Mol. Biol. Rev. **75:**636–663.

Heinig, M., and Frishman, D. (2004) STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. Nucleic Acids Res. **32:**W500–W502.

Holm, L., and Rosenström, P. (2010) Dali server: conservation mapping in 3D. Nucleic Acids Res. **38:**W545–W549.

Kim, H. S., Kim, J., Im, H. N., An, D. R., Lee, M., Hesek, D., Mobashery, S., Kim, J. Y., Cho, K., Yoon, H. J., Han, B. W., Lee, B. I., and Suh, S. W. (2014) Structural basis for the recognition of muramyltripeptide by *Helicobacter pylori* Csd4, a D,L-carboxypeptidase controlling the helical cell shape. Acta Crystallogr. sect. D Biol. Crystallogr. **70**:2800–2812.

Kim, H. S., Im, H. N., An, D. R., Yoon, J. Y., Jang, J. Y., Mobashery, S., Hesek, D., Lee, M., Yoo, J., Cui, M., Choi, S., Kim, C., Lee, N. K., Kim, S. J., Kim, J. Y., Bang, G., Han, B. W., Lee, B. I., Yoon, H. J., and Suh, S. W. (2015) The cell shape-determining Csd6 protein from *Helicobacter pylori* constitutes a new family of L,D-carboxypeptidase. J. Biol. Chem. **290**:25103–25117.

Knott, G. D. (1979) Mlab–a mathematical modeling tool. Comput. Programs Biomed. **10:**271–280.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. **305:**567–580.

Koch, M. K., McHugh, C. A., and Hoiczyk, E. (2011) BacM, an N-terminally processed bactofilin of *Myxococcus xanthus*, is crucial for proper cell shape. Mol. Microbiol. **80**:1031–1051.

Kühn, J., Briegel, A., Mörschel, E., Kahnt, J., Leser, K., Wick, S., Jensen, G. J., and Thanbichler, M. (2010) Bactofilins, a ubiquitous class of cytoskeletal proteins mediating polar localization of a cell wall synthase in *Caulobacter crescentus*. EMBO J. **29**:327–339.

Kusters, J. G., van Vliet, A. H., and Kuipers, E. J. (2006) Pathogenesis of *Helicobacter pylori* infection. Clin. Microbiol. Rev. **19**:449–490.

Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) Computer-aided interpretation of analytical sedimentation data for proteins. In: Analytical Ultracentrifugation in Biochemistry and Polymer Science (Harding, S.E., Rowe, A.J. and Horton, J.C., Eds.), pp. 90–125. Royal Society of Chemistry, Cambridge

Lee, S. Y., Lee, J. H., Chang, H. J., Cho, J. M., Jung, J. W., and Lee, W. (1999) Solution structure of a sweet protein single-chain monellin determined by nuclear magnetic resonance and dynamical simulated annealing calculations. Biochemistry **38:**2340–2346.

Lee, T. K., and Huang, K. C. (2013) The role of hydrolases in bacterial cell-wall growth. Curr. Opin. Microbiol. **16:**760–766.

Lertsethtakarn, P., Ottemann, K. M., and Hendrixson, D. R. (2011) Motility and chemotaxis in *Campylobacter* and *Helicobacter*. Annu. Rev. Microbiol. **65**:389–410.

Malfertheiner, P., Selgrad, M., and Bornschein, J. (2012) *Helicobacter pylori*: clinical management. Curr. Opin. Gastroenterol. **28:**608–614.

Meroueh, S. O., Bencze, K. Z., Hesek, D., Lee, M., Fisher, J. F., Stemmler, T. L., and Mobashery, S. (2006) Three-dimensional structure of the bacterial cell wall peptidoglycan. Proc. Natl. Acad. Sci. USA **103:**4404–4409.

Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. sect. D Biol. Crystallogr. **53:**240–255.

Odintsov, S. G., Sabala, I., Marcyjaniak, M., and Bochtler, M. (2004) Latent LytM at 1.3A resolution. J. Mol. Biol. **335:**775–785.

Otero, L. H., Rojas-Altuve, A., Llarrull, L. I., Carrasco-López, C., Kumarasiri, M., Lastochkin, E., Fishovitz, J., Dawley, M., Hesek, D., Lee, M., Johnson, J. W., Fisher, J. F., Chang, M., Mobashery, S., and Hermoso, J. A. (2013) How allosteric control of *Staphylococcus aureus* penicillin binding protein 2a enables methicillin resistance and physiological function. Proc. Natl. Acad. Sci. USA **110:**16808–16813.

Ottemann, K. M., and Lowenthal, A. C. (2002) Helicobacter pylori uses motility for initial colonization and to attain robust infection. Infect. Immun. **70:**1984–1990.

Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in osillation mode. Methods Enzymol. **276:**307–326.

Ragumani, S., Kumaran, D., Burley, S. K., and Swaminathan, S. (2008) Crystal structure of a putative lysostaphin peptidase from *Vibrio cholerae*. Proteins **72:**1096–1103.

Read, R. J., and Schierbeek, A. J. (1988) A phased translation function. J. Appl. Crystallogr. **21:**490–495.

Roesler, B. M., Rabelo-Gonçalves, E. M., and Zeitune, J. M. (2014) Virulence factors of *Helicobacter pylori*: A review. Clin. Med. Insights Gastroenterol. **7**:9–17.

Roubaud Baudron, C., Franceschi, F., Salles, N., and Gasbarrini, A. (2013) Extragastric diseases and *Helicobacter pylori*. Helicobacter **18**:44–51.

Sabala, I., Jagielska, E., Bardelang, P. T., Czapinska, H., Dahms, S. O., Sharpe, J. A., James, R., Than, M. E., Thomas, N. R., and Bochtler, M. (2014) Crystal structure of the antimicrobial peptidase lysostaphin from *Staphylococcus simulans*. FEBS J. **18:**4112–4122.

Scheffers, D. J., and Pinho, M. G. (2005) Bacterial cell wall synthesis: new insights from localization studies. Microbiol. Mol. Biol. Rev. **69:**585–607.

Schreiber, S., Konradt, M., Groll, C., Scheid, P., Hanauer, G., Werling, H. O., Josenhans, C., and Suerbaum, S. (2004) The spatial orientation of

Helicobacter pylori in the gastric mucus. Proc. Nat. Acad. Sci. USA **101:**5024–5029.

Spencer, J., Murphy, L. M., Conners, R., Sessions, R. B., and Gamblin, S. J. (2010) Crystal structure of the LasA virulence factor from *Pseudomonas aeruginosa*: substrate specificity and mechanism of M23 metallopeptidases. J. Mol. Biol. **396:**908–923.

Sycuro, L. K., Pincus, Z., Gutierrez, K. D., Biboy, J., Stern, C. A., Vollmer, W., and Salama, N. R. (2010) Peptidoglycan crosslinking relaxation promotes *Helicobacter pylori*'s helical shape and stomach colonization. Cell **141:**822–833.

Sycuro, L. K., Rule, C. S., Petersen, T. W., Wyckoff, T. J., Sessler, T., Nagarkar, D. B., Khalid, F., Pincus, Z., Biboy, J., Vollmer, W., and Salama, N. R. (2013) Flow cytometry-based enrichment for cell shape mutants identifies multiple genes that influence *Helicobacter pylori* morphology. Mol. Microbiol. **90**:869–883.

Sycuro, L. K., Wyckoff, T. J., Biboy, J., Born, P., Pincus, Z., Vollmer, W., and Salama, N. R. (2012) Multiple peptidoglycan modification networks modulate *Helicobacter pylori*'s cell shape, motility, and colonization potential. PLoS Path. **8**:e1002603.

Terwilliger, T. C. (2003) Automated main-chain model building by template matching and iterative fragment extension. Acta Crystallogr. sect. D Biol. Crystallogr. **59:**38–44.

Vagin, A., and Teplyakov, A. (2010) Molecular replacement with MOLREP. Acta Crystallogr. sect. D Biol. Crystallogr. **66:**22–25.

Vollmer, W. (2012) Bacterial growth does require peptidoglycan hydrolases. Mol. Microbiol. **86:**1031–1035.

Vollmer, W., and Bertsche, U. (2008) Murein (peptidoglycan) structure, architecture and biosynthesis in *Escherichia coli*. Biochim. Biophys. Acta **1778:**1714–1734.

Vollmer, W., Blanot, D., and De Pedro, M. A. (2008a) Peptidoglycan structure and architecture. FEMS Microbiol. Rev. **32:**149–167.

Vollmer, W., Joris, B., Charlier, P., and Foster, S. (2008b) Bacterial peptidoglycan (murein) hydrolases. FEMS Microbiol. Rev. **32:**259–286.

Wang, X., Yang, X., Yang, C., Wu, Z., Xu, H., and Shen, Y. (2011) Crystal structure of outer membrane protein NMB0315 from *Neisseria meningitidis*. PLoS One **6:**e26845.

Weidel, W., and Pelzer, H. (1964) Bag-shaped macromolecules: A new outlook on bacterial cell walls. Adv. Enzymol. Relat. Areas Mol. Biol. **26**:193–232.

Worku, M. L., Sidebotham, R. L., Walker, M. M., Keshavarz, T., and Karim, Q. N. (1999) The relationship between *Helicobacter pylori* motility, morphology and phase of growth: implications for gastric colonization and pathology. Microbiology **145:** 2803–2811.

Wyckoff, T. J., Taylor, J. A., and Salama, N. R. (2012) Beyond growth: novel functions for bacterial cell wall hydrolases. Trends Microbiol. **20:**540–547.

Yang, D. C., Tan, K., Joachimiak, A., and Bernhardt, T. G. (2012) A conformational switch controls cell wall-remodelling enzymes required for bacterial cell division. Mol. Microbiol. **85:**768–781.

Young, K. D. (2007) Bacterial morphology: why have different shapes? Curr. Opin. Microbiol. **10:**596–600.

Chapter 2

Structural Basis of the Heterodimer Formation between Cell Shape-Determining Proteins Csd1 and Csd2 from *Helicobacter pylori*

2.1. Introduction

Helicobacter pylori is a Gram-negative bacterium that colonizes the stomach of roughly half of the world's population, thus causing a variety of gastrointestinal diseases such as peptic ulcer and gastric cancer (Roesler et al., 2014). It is regarded as a primary factor for gastric cancer development (Roesler et al., 2014) and is the sole Group I carcinogen among pathogenic bacteria according to the classification by the International Agency for Research on Cancer. Recent reports also suggest possible links between *H. pylori* infection and some extra-digestive diseases, including neurodegenerative disorders (Franceschi et al., 2015). Typical treatment regimens for *H. pylori* infection consist of a proton pump inhibitor such as omeprazole and the antibiotics clarithromycin and amoxicillin (or

metronidazole). However, increasing drug resistance requires new therapies and the discovery of new antibiotics (Malfertheiner et al., 2012).

High motility of *H. pylori* is important for its colonization of the human stomach and its survival in its preferred niche, the gastric mucosa (Ottemann and Lowenthal, 2002; Schreiber et al., 2004; Lertsethtakarn et al., 2011). The helical cell shape of *H. pylori* is believed to facilitate efficient colonization of the viscous epithelial mucus layer via a cork-screwing mechanism (Berg and Turner, 1979; Hazell et al., 1986; Worku et al., 1999). It has been shown for helical proteobacteria that a loss of motility and/or helical cell shape causes major defects in the colonization potential of these organisms (Sycuro et al., 2010; Sycuro et al., 2012; Frirdich et al., 2012; Wyckoff et al., 2012).

The peptidoglycan layer of a bacterial cell wall plays a key role not only in protecting cells against the intracellular turgor pressure but also in maintaining the proper cell shape (Scheffers and Pinho, 2005; Vollmer and Bertsche, 2008). It is made of linear polysaccharide chains that consist of alternating β-1,4-linked *N*-acetylglucosamine-*N*-acetylmuramic acid (NAG-NAM) disaccharide units, with a pentapeptide covalently linked to NAM (Vollmer et al., 2008a). In *H. pylori*, the pentapeptide consists of L-Ala^{1}-γ-D-Glu^{2}-*m*DAP^{3}-D-Ala^{4}-D-Ala^{5} (or -Gly^{5}), where *m*DAP refers to meso-2,6-diaminopimelate (Figure 2-1). The neighboring peptides are cross-linked

exclusively by the 4→3 linkage between the α-carboxylic group of D-Ala^4 on one strand (the donor peptide) and the ε-amino group of *m*DAP^3 on another strand (the acceptor peptide) (Weidel and Pelzer, 1964; Costa et al., 1999) to form a mesh-like peptidoglycan (murein) sacculus (Meroueh et al., 2006). In many bacteria, the peptidoglycan layer is remodeled by a number of peptidoglycan hydrolases and by lytic transglycosylases that function in the peptidoglycan maturation, regulation of cell wall growth, cell division, peptidoglycan turnover and recycling, cell lysis and the release of peptidoglycan fragments for host–pathogen interactions (Vollmer et al., 2008b; Vollmer, 2012; Lee and Huang, 2013; Heijenoort, 2011; Lee et al., 2013). In *H. pylori*, cleaving the 4→3 cross-links of peptidoglycan or trimming of peptidoglycan muropeptides affect its helical cell shape. A small number of proteins have been identified to be essential in generating the helical cell shape of *H. pylori* by tailoring the peptidoglycan layer (Wyckoff et al., 2012): an amidase AmiA (Chaput et al., 2006), potential peptidoglycan peptidases Csd1–Csd4 (Sycuro et al., 2010; Sycuro et al., 2012; Bonis et al., 2010; Chan et al., 2015) and Csd6 (Sycuro et al., 2013), and potential regulators Csd5 and CcmA (Sycuro et al., 2010; Sycuro et al., 2012) (Figure 2-2).

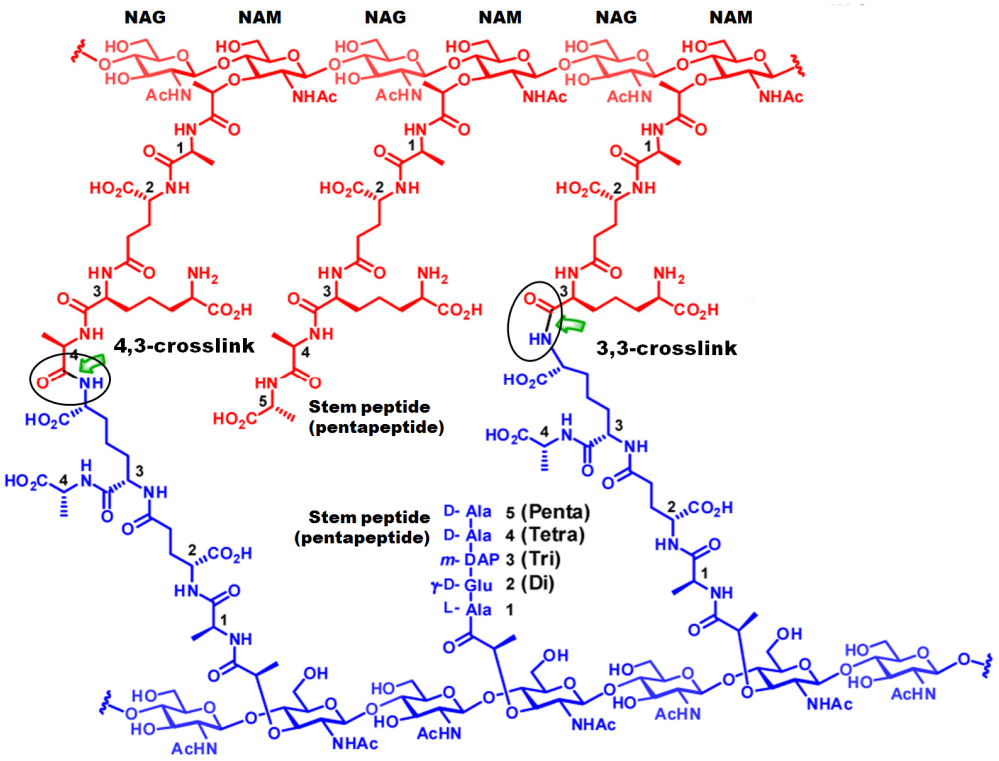


Figure 2-1. Chemical structure of cross-linked peptidoglycan in Gram-negative bacteria (modified from Lee et al., 2015)

The donor strand and the acceptor strand of cross-links are in red and in blue, respectively. Two types of crosslink are indicated by green arrows. In *H. pylori*, peptidoglycan is exclusively cross-linked by the 4,3-crosslinks.

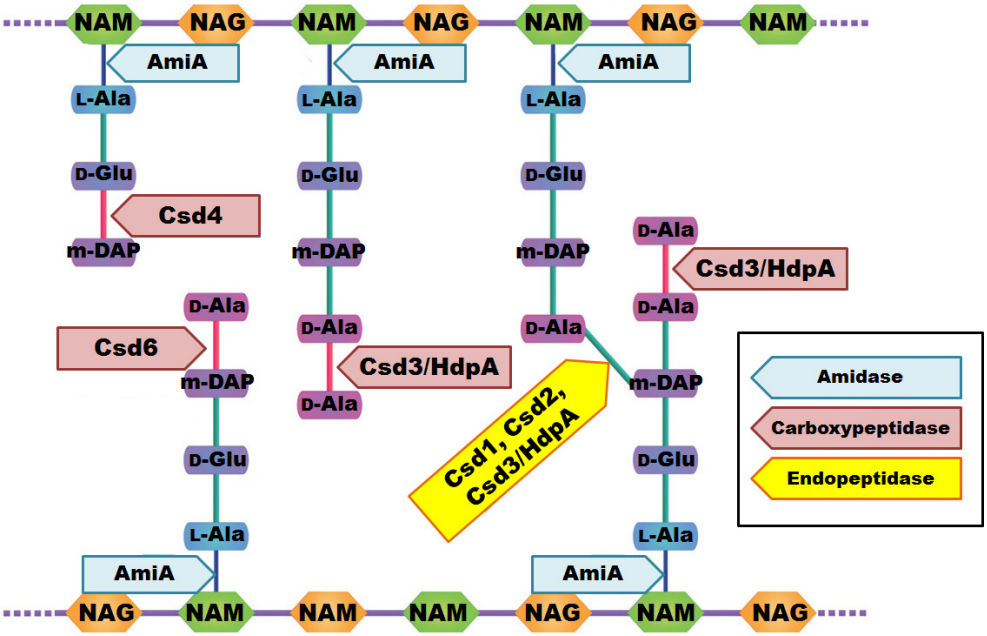


Figure 2-2. Cleavage specificities of peptidoglycan hydrolases in *H. pylori* (modified from Wyckoff et al., 2012)

Peptidoglycan hydrolase class is color-coded as indicated in the black box.

Among them, three Csd proteins (Csd1 and Csd2, as well as Csd3) are M23B metallopeptidase family members and were identified as shape-determining LytM peptidase homologs in *H. pylori*. They all contribute to *H. pylori*'s helical cell shape through alternations in peptidoglycan crosslinking (Sycuro et al., 2010). Deletion mutants of each of three genes resulted in curved rod morphologies, with varying degrees of curvature in Δ*csd3*, as well as showed increased tetra-pentapeptide crosslinking in muropeptide composition (Sycuro et al., 2010). They may catalyze the same reaction as D,D-endopeptidase (D,D-EPase), relaxing the D-Ala^4-*m*DAP^3 peptide bond (4→3 cross-link) in crosslinked dimers of muramyl tetra- and penta-peptides (muramyl-L-Ala^1-γ-D-Glu^2-mDAP^3-D-Ala^4 and muramyl-L-Ala^1-γ-D-Glu^2-mDAP^3-D-Ala^4-D-Ala^5, respectively) (Sycuro et al., 2010). Csd1 shows the broadest conservation with a homolog present in most ε-proteobacteria, whereas Csd2 homologs are found only in *H. pylori* and *H. hepaticus* (Sycuro et al., 2010). Csd3 homologs, which contain an N-terminal extension not present in Csd1 and Csd2, are well conserved throughout the ε-proteobacteria including *H. pylori* (Sycuro et al., 2010). All the Csd1 and Csd3 homologs identified preserve conserved residues in two characteristic motifs (**H**xxx**D** and **H**x**H** motifs), which are predicted to function in catalysis as peptidoglycan endopeptidases or carboxypeptidases (Sycuro et

al., 2010). Crystal structure of N-terminally truncated *H. pylori* Csd3 (HP0506 in 26695 strain) revealed that Csd3 consists of three domains: Domain 1 (residues 42–124), Domain 2 (residues 125–228 and 360–403) and the C-terminal LytM domain (residues 229–359) (An et al., 2015). The LytM domain of Csd3 has a canonical fold with a Zn^{2+}-containing active site. The Zn^{2+} ion is tetra-coordinated with the expected three ligands (His259 and Asp263 of **H**xxx**D** motif and His341 of Hx**H** motif) and one residue (Glu74) from the α3 helix of domain 1. Domain 1 of Csd3 functions as an inhibitory domain by blocking the access of the substrate into the active site in the latent state (An et al., 2015).

In *H. pylori* 26695 strain, the *csd1* gene (HP1543) is located immediately upstream of the *ccmA* gene (HP1542) and downstream of the *csd2* gene (HP1544), resulting in a three-gene locus (Figure 2-3A). Deletion of the *H. pylori ccmA* gene resulted in curved-rod morphology (Figure 2-3B) and showed the increases in tetra-pentapeptide crosslinking in muropeptide composition (Sycuro et al., 2010). Because CcmA does not encode any recognizable peptidase motif, it may function through interaction with Csd peptidases. *H. pylori* Csd1 is 53% similar and 29% identical to Csd2 (Sycuro et al., 2010). They were identified to contain the LytM domain at the C-terminus, but the canonical motifs (HxxxD and HxH motifs) of the LytM

domain were only completed in Csd1. Point mutation of a key histidine residue in Csd1 (H250A), which is one of three metal-ligands conserved in the LytM canonical motifs, resulted in a slightly curved-rod morphology (Figure 2-3B) that is indistinguishable from the null mutant *Δcsd1* (Sycuro et al., 2010).

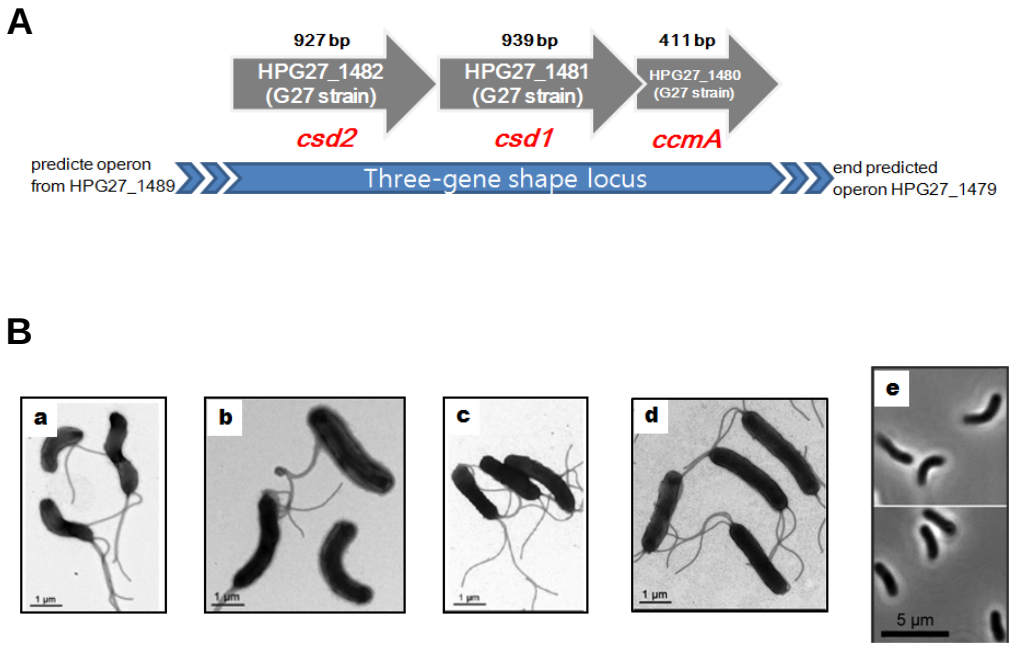


Figure 2-3. *H. pylori* cell shape locus and mutant morphologies (modified from Sycuro et al., 2010)

(A) Three-gene shape locus in *H. pylori* G27 strain. (B) Transmission electron microscopy (TEM) images of wild-type (a), *Δcsd2* (b), *Δcsd1* (c), *ΔccmA* (d), and Csd1 H250A point mutant (e).

To provide insight into the molecular functions of *H. pylori* Csd1 and Csd2 proteins, I have carried out their structural characterizations. I have discovered that (i) Csd2 exists in monomer-dimer equilibrium in solution by equilibrium sedimentation and (ii) Csd1 and Csd2 form a stable 1:1 heterodimer. I have determined the crystal structures of the $Csd2_{121–308}$ homodimer and the heterodimer between $Csd1_{125–312}$ and $Csd2_{121–308}$. The overall structures of $Csd1_{125–312}$ and $Csd2_{121–308}$ monomers are similar to each other, consisting of a helical domain and a LytM domain. The helical domains of both Csd1 and Csd2 play a key role in the formation of homodimers or heterodimers. LytM domains of Csd1 and Csd2 share the same overall fold but a functionally significant difference exists in their active sites. The Csd1 LytM domain contains a catalytic site with a Zn^{2+} ion, which is coordinated by three conserved ligands and two water molecules, whereas Csd2 has a degenerate LytM domain with incomplete metal ligands and no metal ion is bound. I also observed two types of non-canonical Zn^{2+}-coordination in the active site of Csd1 LytM domain. In one $Csd2_{121–308}$ chain of the heterodimer models between $Csd1_{125–312}$ and $Csd2_{121–308}$, the C-terminal tail of Csd2 is bound to the central groove of the Csd1 LytM domain and defines the substrate binding site. The structural knowledge

from this work sheds light on the events that regulate the cell wall in *H. pylori*.

2.2. Material and methods

2.2.1. Cloning, expression, and purification of Csd2

Five different constructs of Csd2 (residues 54–308, 63–308, 77–308, 121–308, and 145–251) were individually expressed in a soluble form (Table 2-1). Among them, crystals were produced from the $Csd2_{121–308}$ construct only.

For crystal structure of $Csd2_{121–308}$, the *csd2* gene (*hp1544* from *H. pylori* 26695 strain) was PCR-amplified and was cloned into the expression vector pET-28b(+) (Novagen) using the NdeI and XhoI restriction enzyme sites. The resulting recombinant $Csd2_{121–308}$ is fused with hexahistidine-containing tags at both N- and C-termini (MGSSHHHHHHSSGLVPRGSH at the N-terminus and LEHHHHHH at the C-terminus). The recombinant protein was overexpressed in *Escherichia coli* Rosetta 2(DE3)pLysS cells, using the Luria Broth culture medium. Protein expression was induced by 0.5 mM isopropyl β-D-thiogalactopyranoside and the cells were incubated for additional 15 h at 30°C following growth to mid-log phase at 37°C. The cells were harvested and were suspended in an ice-cold lysis buffer [20 mM Tris-HCl at pH 7.9, 500 mM sodium chloride, 50 mM imidazole, and 10%

(*v/v*) glycerol] containing 1 mM phenylmethylsulfonyl fluoride. The cell was lysed by sonication. After centrifugation at 36,000 g for 1 h at 4°C, the cell debris was discarded and the supernatant was applied to an affinity chromatography column of HiTrap Chelating HP (GE Healthcare), which was previously equilibrated with the lysis buffer. The column was eluted with a linear gradient from 25 to 500 mM imidazole. The recombinant $Csd2_{121–308}$ protein eluted at 150–200 mM imidazole concentration. The eluted protein was further purified by gel filtration on a HiLoad 16/60 Superdex 200 prep-grade column (GE Healthcare), which was previously equilibrated with 20 mM HEPES at pH 7.5 and 200 mM sodium chloride. Peak fractions containing the $Csd2_{121–308}$ protein were pooled and concentrated to 10 mg ml^{-1} (0.40 mM) for crystallization.

To perform SEC-MALS (size-exclusion chromatography with multi-angle static light scattering) and equilibrium sedimentation experiments, I also PCR-amplified the *csd2* gene covering residues 145–251 and cloned it into the expression vector pET-21a(+) (Novagen) to express the recombinant $Csd2_{145–251}$ protein fused with a hexahistidine-containing tag (LEHHHHHH) at the C-terminus. Identical procedures were followed for protein expression and purification of both $Csd2_{121–308}$ and $Csd2_{145–251}$. The recombinant

$Csd2_{145–251}$ protein eluted at 120–150 mM imidazole concentrations during affinity chromatography on a HiTrap Chelating HP column (GE Healthcare). The eluted protein was further purified by gel filtration on a HiLoad 16/60 Superdex 200 prep-grade column (GE Healthcare), which was previously equilibrated with 20 mM HEPES at pH 7.5 and 200 mM sodium chloride.

2.2.2. Cloning, expression, and purification of Csd1-Csd2 complexes

I initially tried to express and purify the Csd1 protein alone using four different constructs (Table 2-2). The construct covering the residues 125–312 of *H. pylori* Csd1 (HP1543 from 26695 strain) was cloned into the expression vector pET-28b(+) (Novagen) using the NdeI and XhoI restriction enzyme sites. This construct contains the hexahistidine-containing tags at both N- and C-termini. Three other constructs were also PCR-amplified and were cloned into the expression vector pET-21a(+), resulting in the recombinant $Csd1_{54–312}$ (residues 54–312), $Csd1_{75–312}$ (residues 75–312), and $Csd1_{91–312}$ (residues 91–312) proteins, which are fused with the hexa-histidine containing tag at the C-terminus. All recombinant Csd1 proteins were overexpressed in *E. coli* Rosetta 2(DE3)pLysS cells using the

Luria Broth culture medium. Protein expression was induced by 0.5 mM isopropyl β-D-thiogalactopyranoside and the cells were incubated for additional 15 h at 30°C following growth to mid-log phase at 37°C. All of the above Csd1 constructs were expressed in an insoluble form despite extensive screening of the cell culture condition.

To test the possible complex formation between Csd1 and Csd2 proteins (Table 2-3), the $Csd2_{121–308}$ construct was cloned into the expression vector pET-21a(+) to express the recombinant $Csd2_{121–308}$ protein without a hexahistidine- containing tag at both N- or C-termini. Cells expressing each of the above four Csd1 constructs with fusion tag(s) and $Csd2_{121–308}$ without a fusion tag were grown separately. Cells were mixed in an approximate mass ratio of 3:1 for Csd1 and Csd2 to account for different expression levels. The mixed cells were diluted in the ice-cold lysis buffer containing 1 mM phenylmethylsulfonyl fluoride and lysed using sonication. After centrifugation at 36,000 g for 1 h at 4°C to discard the cell debris, the supernatant was applied to the affinity chromatography column of HiTrap Chelating HP (GE Healthcare), which was previously equilibrated with the lysis buffer. The column was eluted with a linear gradient from 25 to 500 mM imidazole. The complexes between each of the above four Csd1

constructs and $Csd2_{121–308}$ were eluted at 150–250 mM imidazole concentration. The complex formation was confirmed by SDS-PAGE. I further confirmed by SEC-MALS that $Csd1_{54–312}$ and $Csd2_{121–308}$ form a heterodimer in solution. For crystallization, the complex between $Csd1_{125–312}$ and $Csd2_{121–308}$ was further purified by gel filtration on a HiLoad 16/60 Superdex 200 prep-grade column (GE Healthcare), which was previously equilibrated with 20 mM HEPES at pH 7.5 and 200 mM sodium chloride. Peak fractions containing the $Csd1_{125–312}$ and $Csd2_{121–308}$ complex were pooled and concentrated to 8 mg ml^{-1} (0.15 mM heterodimer concentration) for crystallization.

Table 2-1. Construct design and protein expression for *H. pylori* Csd2

Construct	*Fusion tag*	*Protein expression*	*Crystallization*	*Other experiements*
Csd2 (54–308)	C-term. tag	Soluble	X	-
Csd2 (63–308)	C-term. tag	Soluble	X	-
Csd2 (77–308)	C-term. tag	Soluble	X	-
Csd2 (121–308)	N,C-term. tags	Soluble	O	Equilibrium sedimentation & SEC-MALS
Csd2 (145–251)	C-term. tag	Soluble	X	Equilibrium sedimentation & SEC-MALS

Table 2-2. Construct design and protein expression for *H. pylori* Csd1

Construct	*Fusion tag*	*Protein expression*
Csd1 (54–312)	C-term. tag	Insoluble
Csd1 (75–312)	C-term. tag	Insoluble
Csd1 (91–312)	C-term. tag	Insoluble
Csd1 (125–312)	N,C-term. tags	Insoluble

Table 2-3. Construct design and protein expression for Csd1-Csd2 complex

Construct	*Fusion tag*	*Complex formation*	*Crystallization*	*Experiement*
Csd1 (54–312)	C-term. tag	$Csd1_{54–312}$- $Csd2_{121–308}$	X	SEC-MALS
Csd2 (121–308)	no-tag			
Csd1 (75–312)	C-term. tag	$Csd1_{75–312}$- $Csd2_{121–308}$	X	-
Csd2 (121–308)	no-tag			
Csd1 (91–312)	C-term. tag	$Csd1_{91–312}$- $Csd2_{121–308}$	X	-
Csd2 (121–308)	no-tag			
Csd1 (125–312)	N,C-term. tags	$Csd1_{125–312}$- $Csd2_{121–308}$	O	
Csd2 (121–308)	no-tag			

2.2.3. Crystallization and X-ray data collection

Crystals of $Csd2_{121–308}$ were grown at 23°C by the sitting-drop vapor diffusion method using the Mosquito robotic system (TTP Labtech). Each sitting drop (0.4 μl) was prepared by mixing equal volumes of the protein solution and the reservoir solution. The sitting drop was equilibrated against 100 μl of the reservoir solution. Rod-shaped crystals (Figure 2-4) were obtained with the reservoir solution of 100 mM HEPES at pH 7.0, and 30% (*w/v*) Jeffamine ED-2001. Crystals were cryoprotected in the reservoir solution supplemented with 10% (*v/v*) glycerol, and were flash-frozen in a nitrogen gas stream at 100 K. Native data were collected to 1.80 Å resolution using the ADSC Q315r CCD detector at the beamline BL-5C of Pohang Light Source, Pohang, Korea (Table 2-4). Raw X-ray diffraction data were processed and scaled using the program suit HKL2000 (Otwinowski and Minor, 1997). They belong to the space group $P6_122$, with unit cell parameters of $a = b = 140.8$ Å, and $c = 40.2$ Å. Assuming the presence of one $Csd2_{121–308}$ protein in the asymmetric unit, the Mattew's coefficient and solvent content are 2.80 $Å^3\ Da^{-1}$ and 56.1 %, respectively.

Crystals of the heterodimer between $Csd1_{125–312}$ and $Csd2_{121–308}$ were grown at 23°C by the sitting-drop vapor diffusion method. Each sitting drop

prepared by mixing 0.2 μl each of protein solution and the reservoir solution. The sitting drop was equilibrated against 100 μl of reservoir solution. Clusters of needle-like crystals (Figure 2-4) were obtained with reservoir solution of 100 mM Tris-HCl at pH 8.5, and 25% (*w/v*) PEG3350. I initially tried to optimize this crystallization condition but failed to improve the crystal quality. Therefore, microseeding was employed. A stock solution of microseed crystals was prepared by crushing the needle-like crystals in 50 μl of the reservoir solution using the Seed Bead kit (Hampton Research) and by diluting hundred-fold with the reservoir solution. Clusters of rod-shaped crystals (Figure 2-4) were obtained when I used a reservoir solution consisting of 100 mM HEPES at pH 7.0, and 30% (*w/v*) Jeffamine ED-2001 and 4 μl of the sitting drop, which was prepared by mixing 2 μl of the protein solution, 1.6 μl of the reservoir solution, and 0.4 μl of the microseed crystal stock solution. They were cryoprotected in the reservoir solution supplemented with 10% (*v/v*) glycerol and were flash-frozen in a nitrogen gas stream at 100 K. Two sets of native data were collected from different crystals to 2.27 Å and 2.35 Å resolutions using the ADSC Q270 CCD detector at the beamline BL-7A of Pohang Light Source, Pohang, Korea. Raw X-ray diffraction data were processed and scaled using the program suit HKL2000 (Otwinowski and Minor, 1997). They belong to the space group

$P2_1$, with unit cell parameters of a = 53.4 Å, b = 82.4 Å, c = 76.3 Å, and β = 106.0 Å for Csd1-Csd2 dimer I data and a = 53.1 Å, b = 80.0 Å, c = 74.6 Å, and β = 104.4 Å for Csd1-Csd2 dimer II data. Data collection statistics are summarized in Table 2-4.

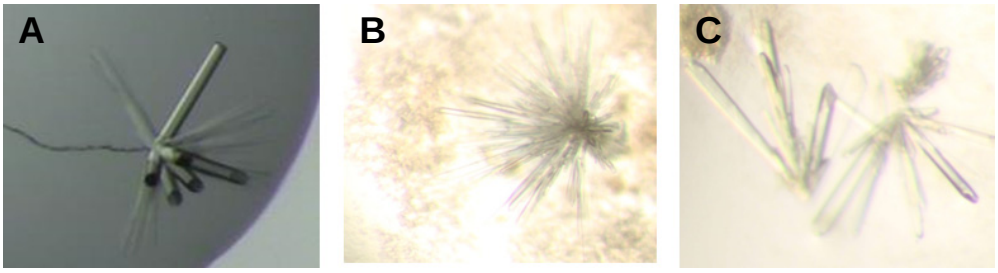


Figure 2-4. C**rystals of Csd2$_{121–308}$ and the heterodimer between Csd1$_{125–312}$ and Csd2$_{121–308}$**

(A) Csd2$_{121–308}$ homodimer crystals. (B) Initial crystals of the heterodimer between Csd1$_{125–312}$ and Csd2$_{121–308}$. (C) Refined crystals of the heterodimer between Csd1$_{125–312}$ and Csd2$_{121–308}$.

Table 2-4. Statistics of data collection

Data collection			
Data set	Csd2-Csd2 dimer	Csd1-Csd2 dimer I	Csd1-Csd2 dimer II
Space group	$P6_122$	$P2_1$	$P2_1$
Unit cell lengths, *a, b, c* (Å)	140.8, 140.8, 40.2	53.4, 82.4, 76.3	53.1, 80.0, 74.6
Unit cell angles, α, β, γ (°)	90, 90, 120	90, 106.0, 90	90, 104.4, 90
X-ray wavelength (Å)	0.97960	0.97933	0.97933
Resolution range (Å)	50.0–1.80 (1.83–1.80)[a]	50.0–2.27 (2.31–2.27)[a]	50.0–2.35 (2.39–2.35)[a]
Total / unique reflections	365,382 / 22,028	118,925 / 29,291	155,055 / 25,521
Completeness (%)	99.8 (98.6)[a]	99.8 (100.0)[a]	99.9 (100.0)[a]
$\langle I \rangle / \langle \sigma_I \rangle$	47.1 (6.9)[a]	20.8 (4.0)[a]	26.6 (3.9)[a]
R_{merge}[b] (%)	12.8 (75.1)[a]	13.5 (73.0)[a]	13.7 (88.6)[a]
R_{rim}[c] (%)	13.2 (77.4)[a]	15.5 (83.6)[a]	15.0 (96.5)[a]
R_{pim}[d] (%)	3.2 (18.6)[a]	7.6 (40.3)[a]	6.0 (37.9)[a]
$CC_{1/2}$[e] (%)	99.8 (94.1)[a]	99.3 (70.1)[a]	99.5 (75.9)[a]

Footnotes to Table 2-4

[a] Values in parentheses refer to the highest resolution shell.

[b] $R_{merge} = \Sigma_h\Sigma_i \mid I(h)_i - \langle I(h)\rangle \mid / \Sigma_h\Sigma_i I(h)_i$, where $I(h)$ is the intensity of reflection h, Σ_h is the sum over all reflections, and Σ_i is the sum over i measurements of reflection h.

[c] $R_{rim} = \Sigma_{hkl}\{N(hkl) / [N(hkl) - 1]\}^{1/2} \Sigma_i \mid I_i(hkl) - \langle I(hkl)\rangle \mid / \Sigma_{hkl}\Sigma_i I_i(hkl)$. The redundancy-independent merging R factor gives the precision of individual intensity (Diederichs and Karplus, 1997).

[d] $R_{pim} = \Sigma_{hkl}\{1 / [N(hkl) - 1]\}^{1/2} \Sigma_i \mid I_i(hkl) - \langle I(hkl)\rangle \mid / \Sigma_{hkl}\Sigma_i I_i(hkl)$. The precision indicating merging R factor describes the precision of the averaged intensity (Weiss, 2001).

[e] $CC_{1/2}$ is the correlation coefficient of the mean intensities between two random half-sets of data.

2.2.4. Structure determination and refinement

The structure of $Csd2_{121–308}$ homodimer was determined by molecular replacement utilizing the program MOLREP (Vagin and Teplyakov, 2010), with domain 3 of an outer-membrane protein NMB0315 from *Neisseria meningitidis* (PDB code 3SLU) as a search model. Domain 3 of NMB0315 shows 24% sequence identity with residues Lys149–Asp259 of $Csd2_{121–308}$. Structures of the heterodimer between $Csd1_{125–312}$ and $Csd2_{121–308}$ were solved by molecular replacement using the refined monomer model of $Csd2_{121–308}$ as a search model. The sequence identity between $Csd1_{125–312}$ and $Csd2_{121–308}$ is 39%. Manual model building was done using the program COOT (Emsley et al., 2010) and the models were refined with the programs REFMAC5 (Murshudov et al., 1997), including the bulk solvent correction. A total of 5% of the data was randomly set aside as test data for the calculation of R_{free} (Brünger, 1992). The stereochemistry of the refined models was assessed by MolProbity (Chen et al., 2010). Refinement statistics are summarized in Table 2-5. The coordinates and structure factors have been deposited in the Protein Data Bank under accession codes 5J1K for Csd2-Csd2 dimer of $Csd2_{121–308}$ homodimer and 5J1L and 5J1M for Csd1-Csd2 dimer I and dimer II of the $Csd1_{125–312}$-$Csd2_{121–308}$ complex.

Table 2-5. Refinement statistics

Data set	Csd2-Csd2 dimer	Csd1-Csd2 dimer I	Csd1-Csd2 dimer II
Resolution range (Å)	50.0–1.80	50.0–2.27	50.0–2.35
Rwork / Rfree[a] (%)	18.1 / 21.8	17.4 / 23.3	16.8 / 23.3
No. of non-hydrogen atoms / average *B*-factor (Å^2)			
Protein	1,457 / 22.5	5,432 / 32.8	5,418 / 45.0
Water oxygen	97 / 26.8	160 / 30.0	140 / 40.0
Glycerol	12 / 51.5	-	-
Zn^{2+} ion	-	2 / 63.0	2 / 59.3
R.m.s. deviations from ideal geometry			
Bond lengths (Å)	0.011	0.010	0.010
Bond angles (°)	1.43	1.47	1.40
Ramachandran[b]			
Favored / Outliers (%)	97.2 / 0.0	95.2 / 0.0	95.1 / 0.0
Poor rotamers (%)	0.63	0.68	0.68

Footnotes to Table 2-5

[a] $R_{work} = \Sigma \mid |F_{obs}| - |F_{calc}| \mid / \Sigma |F_{obs}|$, where R_{free} is calculated for a randomly chosen 5% of reflections, which were not used for structure refinement and R_{work} is calculated for the remaining reflections.

[b] Values obtained using *MolProbity*.

2.2.5. Identification of Zn^{2+} binding by anomalous diffraction data

To confirm the identity of a metal ion bound to the catalytic site of Csd1, a set of single-wavelength anomalous diffraction data was collected at 100 K from a crystal of the complex between $Csd1_{125–312}$ and $Csd2_{121–308}$ using the X-ray wavelength of 1.2826 Å (Zn^{2+} absorption edge) at the beamline 7A of Pohang Light Source. Raw data were processed and scaled using HKL2000 (Table 2-6). Anomalous difference maps were calculated using the program FFT of the CCP4i software package (Read and Schierbeek, 1988).

Table 2-6. Statistics for data collection of Zn SAD data set

Data collection	
Data set	Zn SAD (Csd1-Csd2 dimer)
Space group	$P2_1$
Unit cell lengths, a, b, c (Å)	53.5, 80.4, 74.6
Unit cell angles, α, β, γ (°)	90, 104.7, 90
X-ray wavelength (Å)	1.2826
Resolution range (Å)	50.0–2.90 (2.95–2.90)[a]
Total / unique reflections	122,620 / 13,536[b]
Completeness (%)	98.6 (87.6)[a,b]
$\langle I \rangle / \langle \sigma_I \rangle$	49.7 (21.5)[a,b]
R_{merge}[c] (%)	16.7 (72.1)[a,b]
R_{rim}[d] (%)	17.8 (76.1)[a,b]
R_{pim}[e] (%)	6.0 (24.2)[a,b]
$CC_{1/2}$[f] (%)	99.3 (95.4)[a,b]

Footnotes to Table 2-6

[a] Values in parentheses refer to the highest resolution shell.

[b] Friedel pairs were treated as separate observations.

[c] $R_{merge} = \Sigma_h \Sigma_i | I(h)_i - < I(h) > | / \Sigma_h \Sigma_i I(h)_i$, where $I(h)$ is the intensity of reflection h, Σ_h is the sum over all reflections, and Σ_i is the sum over i measurements of reflection h.

[d] $R_{rim} = \Sigma_{hkl}\{N(hkl) / [N(hkl) - 1]\}^{1/2} \Sigma_i | I_i(hkl) - <I(hkl)> | / \Sigma_{hkl}\Sigma_i I_i(hkl)$. The redundancy-independent merging R factor gives the precision of individual intensity (Diederichs and Karplus, 1997).

[e] $R_{pim} = \Sigma_{hkl}\{1 / [N(hkl) - 1]\}^{1/2} \Sigma_i | I_i(hkl) - <I(hkl)> | / \Sigma_{hkl}\Sigma_i I_i(hkl)$. The precision indicating merging R factor describes the precision of the averaged intensity (Weiss, 2001).

[f] $CC_{1/2}$ is the correlation coefficient of the mean intensities between two random half-set of data.

2.2.6. SEC-MALS

SEC-MALS experiments were performed at 23°C using an HPLC system that was connected with a MALS detector (DAWN HELEOS-II, Wyatt Technology) and a differential refractive index detector (Optilab T-Rex, Wyatt Technology). The samples were two Csd2 constructs [$Csd2_{121–308}$ fused with both N- and C-terminal tags and $Csd2_{145–251}$ fused with a C-terminal tag], and a complex between $Csd1_{54–312}$ (fused with a C-terminal tag) and $Csd2_{121–308}$ (without a fusion tag). A size-exclusion chromatography column (WTC-015S5, Wyatt Technology) was pre-equilibrated with the buffer (20 mM HEPES at pH 7.5 and 300 mM sodium chloride) at a flow rate of 0.5 ml min^{-1} and was calibrated using a bovine serum albumin protein standard. Protein samples were injected at a flow rate of 0.5 ml min^{-1}. Weight-averaged molar masses were calculated from the elution data using the ASTRA software (Wyatt Technology).

2.2.7. Equilibrium sedimentation

Equilibrium sedimentation experiments were performed in six-sector cells using a Beckman ProteomeLab XL-A analytical ultracentrifuge for $Csd2_{121–308}$ (fused with both N- and C-terminal tags) and $Csd2_{145–251}$ (fused

with a C-terminal tag) in 20 mM HEPES at pH 7.5 and 200 mM sodium chloride at 4°C. The protein samples were measured at two different speeds (30,000 and 35,000 rpms) and two different protein concentrations (3.5 and 5.1 μM for $Csd2_{121–308}$, and 10.1 and 14.5 μM for $Csd2_{145–251}$, respectively). The protein concentrations were calculated using ε_{280nm} = 25,440 and 8,940 M^{-1} cm^{-1} for $Csd2_{121–308}$ and $Csd2_{145–251}$, respectively. All measured data fit well to a reversible monomer-dimer (1x-2x) equilibrium model for both $Csd2_{121–308}$ and $Csd2_{145–251}$. Representative results for $Csd2_{121–308}$ (measured at 35,000 rpm and 5.1 μM protein concentration) and $Csd2_{145–251}$ (measured at 35,000 rpm and 14.5 μM protein concentration) are presented. The monomer-dimer (1x-2x) equilibrium model for $Csd2_{121–308}$ gave the weighted root-mean-square (r.m.s.) error value of 8.52×10^{-3} with the Ka value of 2.03×10^{5} M^{-1}. The monomer-dimer (1x-2x) equilibrium model for $Csd2_{140–251}$ gave the r.m.s. error value of 9.60×10^{-3} with the Ka value of 2.20×10^{4} M^{-1}.

2.3. Results and discussion

2.3.1. $Csd2_{121–308}$ forms a dimer in the crystal

The crystal structure of $Csd2_{121–308}$ was determined using a hexagonal crystal ('Csd2-Csd2 dimer' in Table 2-4) with one monomer in the asymmetric unit. The model of $Csd2_{121–308}$ monomer accounting for residues His120–Asp301 was refined at 1.80 Å to R_{work} and R_{free} values of 18.1% and 21.8%, respectively (Table 2-5). It consists of three α-helices, two 3_{10}-helices, and ten β-strands, which are arranged in the order of α1-β1-η1-β2-β3-β4-β5-β6-β7-β8-β9-β10-α2-η2-α3 (Figure 2-5A). The $Csd2_{121–308}$ monomer can be divided into two structural domains: the helical domain (residues His120–Gly139 and Asp259–Asp301) and the LytM domain (residues Met140–Leu258). The helical domain is made of four helices (α1, α2, η2, and α3) and the LytM domain is inserted between helices α1 and α2 of the helical domain. The LytM domain of $Csd2_{121–308}$ adopts the canonical fold, in which a central seven-stranded anti-parallel β-sheet (β1↑-β2↓-β9↑-β6↓-β5↑-β4↓-β7↑) forms a two-layered sandwich with a smaller anti-parallel β-sheet (β3↓-β8↑-β6↓) (Figure 2-5B). The long, highly curved strand β6 is shared between these two sheets. A short strand β10 forms a mini, anti-parallel β-

sheet (β9↑-β10↓) with the C-terminal side of β9 of the central β-sheet. A DALI search with the Csd2 LytM domain reveals that it exhibits high structural similarity with other LytM domains of the M23 peptidase family, with Z-scores up to 16.7 (Table 2-7). Essentially identical results were obtained when the whole chain of $Csd2_{121–308}$ was used for the DALI search. It appears that the helical domain of $Csd2_{121–308}$ is structurally unique.

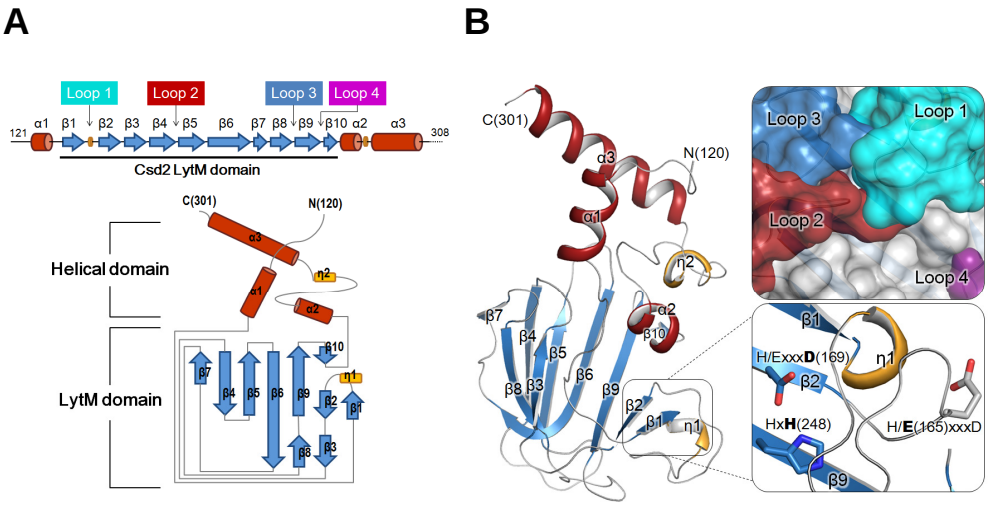


Figure 2-5. Overall monomer structure of *H. pylori* Csd2$_{121–308}$

(A) Schematic representation of secondary structures of $Csd2_{121–308}$ and topology diagram of $Csd2_{121–308}$. Secondary structures have been defined by the STRIDE program (Heinig and Frishman, 2004). α-Helices, β-strands, 3_{10}-helices, and loops are shown as cylinders (colored in red), arrows (skyblue), flat cylinders (yellow), and solid lines (grey), respectively. Loop 1 (β1-β2 loop; cyan), Loop 2 (β4-β5 loop; red), Loop 3 (β8-β9 loop; skyblue), and Loop 4 (β9-β10 loop; purple) form the putative substrate-binding groove of the Csd2 LytM domain. A dotted line indicates the disordered C-terminal region. (B) Ribbon diagram of $Csd2_{121–308}$ monomer structure (chain A of Csd2-Csd2 dimer), colored as in the topology diagram in A. Close-up views on the right represent the surface representation of the putative substrate-binding groove formed by four loops of the LytM domain (top) and the degenerated active site without a metal ion (bottom). Close-up views on the right have different orientations from the monomer ribbon diagram on the left to show the details more clearly. Side chains of Glu165, Asp169, and His248, corresponding to the conserved Zn^{2+}-coordinating residues, are shown in stick models. Asp169 and His248 belong to the Hxxx**D** and Hx**H** motifs of LytM domains, respectively. In Csd2, Glu165 replaces the histidine residue in the HxxxD motif and it is indicated by the modified H/ExxxD motif. No Zn^{2+} ion is bound to the Csd2 LytM domain.

Table 2-7. Structural similarity searches with the Csd2 LytM domain

	Protein name		PDB code	R.m.s. deviation	Z-score	Sequence identity
Csd2 LytM domain	Three-domain protein	Lysostaphin peptidase (VC0503) from *Vibrio cholera*	2GU1	1.7 Å (107 Cα)	16.7	32%
		Outer-membrane protein (NMB0315) from *Neisseria meningitidis*	3SLU	1.9 Å (106 Cα)	16.0	39%
		Cell shape determinant 3 (Csd3; HP0506) from *Helicobacter pylori*	4RNY	1.8 Å (102 Cα)	15.5	28%
	Pro-protein	Glycylglycine endopeptidase (LtyM) from *Staphylococcus aureus*	2B0P	2.2 Å (110 Cα)	15.6	28%
		Lysostaphin from *Staphylococcus simulans*	4QP5	2.0 Å (111 Cα)	15.7	26%
	Non-peptidase protein	Murein hydrolase activator (EnvC) from *Escherichia coli*	4BH5	1.8 Å (107 Cα)	16.1	31%
		Stage II sporulation protein Q (SpollQ) from *Bacillus subtilis*	3UZ0	2.0 Å (113 Cα)	14.6	31%

Unexpectedly, the $Csd2_{121–308}$ monomer (referred to as chain A) in the asymmetric unit of the crystal forms a tight symmetric side-by-side homodimer with a neighboring $Csd2_{121–308}$ monomer (chain A') from an adjacent asymmetric unit (Figure 2-6A). In this homodimer of $Csd2_{121–308}$, a surface area of 1,550 $Å^2$ per monomer is buried at the interface, as analyzed by the PISA sever (Krissinel and Henrick, 2007). The pair of helices α1 and α3 from one helical domain of $Csd2_{121–308}$ pack against another pair of helices (α1' and α3') from the helical domain of another $Csd2_{121–308}$ in the adjacent asymmetric unit to form a four-helix bundle through hydrophobic and hydrogen bond interactions around a crystallographic two-fold symmetry axis (Figure 2-6A). At the interface between helix pairs, the side chain of Gln297 forms hydrogen bonds with the main chain oxygen of Val294 and the side chain of Glu298 (Figure 2-6B). Numerous hydrophobic side chains are present at the interface, which are distributed in both the helix pair (α1 and α3) and the LytM domain (β4, β5-β6 loop, β6-β7 loop, β7-β8 loop, and β9-β10 loop) (Figure 2-6C).

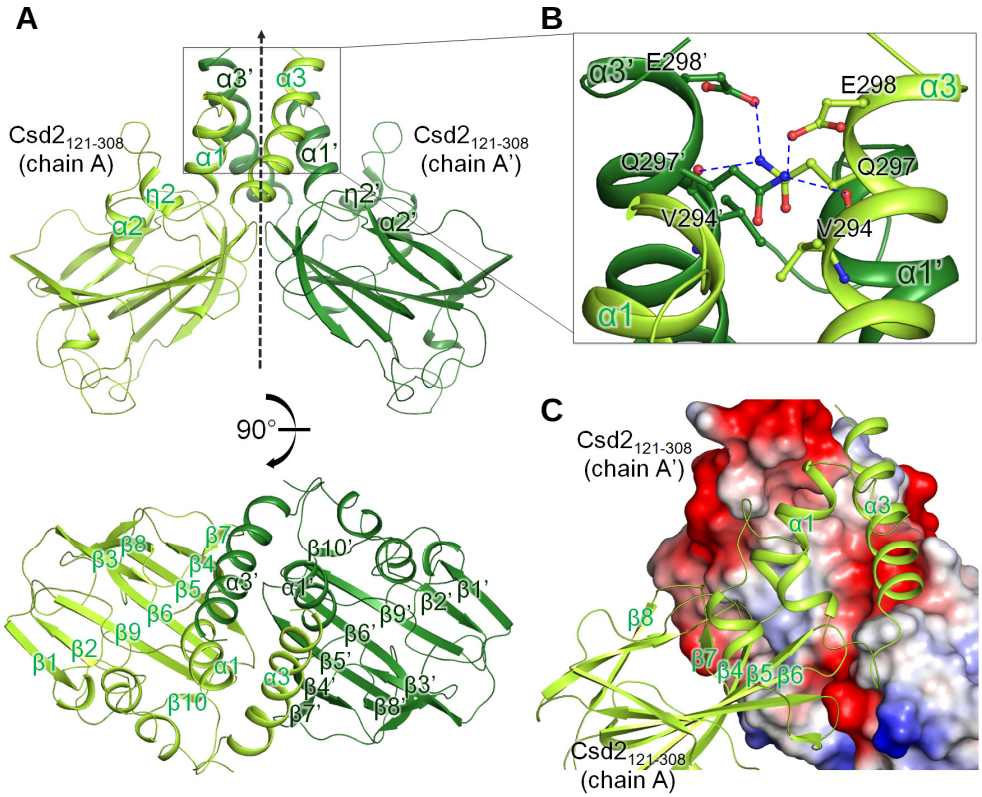


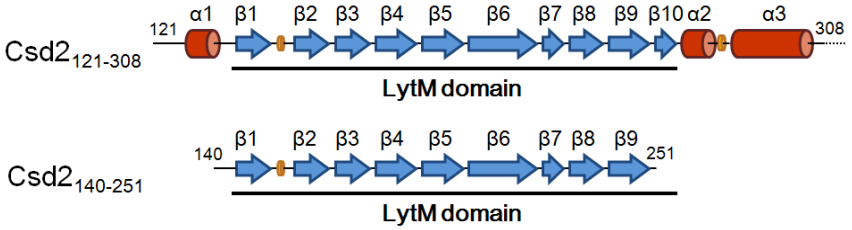
Figure 2-6. *H. pylori* Csd2$_{121–308}$ exists as a homodimer in the crystal

(A) Two different views of the *H. pylori* $Csd2_{121–308}$ homodimer structure are shown in ribbon diagram. One $Csd2_{121–308}$ monomer (chain A colored in yellow-green) and the other $Csd2_{121–308}$ monomer (chain A' colored in darker green) from the adjacent asymmetric unit form a homodimer around a crystallographic two-fold symmetry axis (indicated by a dotted arrow) in the crystal. The secondary structure elements of the helical domain are labeled in the side view (top), while most of the secondary structures are labeled in the top view (bottom). (B) Close-up view of the dimer interface. Residues at the dimer interface are shown in stick models. Blue dotted lines represent hydrogen bonds. (C) One $Csd2_{121–308}$ monomer (chain A) is shown as a ribbon diagram (in yellow-green) and the other $Csd2_{121–308}$ monomer (chain A' from the adjacent asymmetric unit) is shown in the electrostatic surface diagram. This view of the dimer is slightly different from the top view in (A) to show the details more clearly. The dimer interface is hydrophobic in the center and is surrounded by negatively charged surfaces.

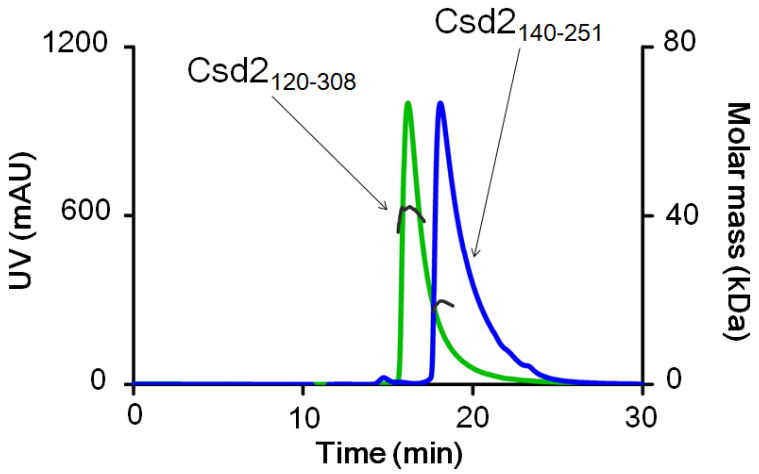
2.3.2. Csd2 exists as monomer-dimer equilibrium in solution

Since $Csd2_{121–308}$ was found to exist as a homodimer in the crystal, I investigated the oligomeric state of Csd2 (using $Csd2_{121–308}$ and $Csd2_{140–251}$) in solution by both SEC-MALS and equilibrium sedimentation experiments. The shorter $Csd2_{140–251}$ construct covers most of the LytM domain except the short β10 strand and lacks the entire helical domain (Figure 2-7A). Molecular masses estimated by SEC-MALS are 41.2 and 19.3 kDa, respectively, for $Csd2_{121–308}$ (with the calculated molecular mass of 24.4 kDa, including the N-terminal and C-terminal fusion tags) and $Csd2_{140–251}$ (with the calculated molecular mass of 13.3 kDa, including the C-terminal fusion tag) (Figure 2-7B). The measured masses are larger than the calculated mass of monomeric species but are smaller than the calculated mass of dimeric species for both constructs of Csd2, making it difficult to assign unambiguously the oligomeric state of these Csd2 proteins in solution.

A



B



	Calculated (kDa)	MALS (kDa)
$Csd2_{120-308}$	24.4	41.2 (±0.36 %)
$Csd2_{140-251}$	13.3	19.3 (±0.27 %)

Figure 2-7. SEC-MALS to determine the oligomeric state of *H. pylori* Csd2 in solution

(A) Two Csd2 constructs ($Csd2_{121–308}$ and $Csd2_{140–251}$) used in these experiments are schematically represented with the secondary structure elements colored as in Figure 2-5A. $Csd2_{140–251}$ lacks the helical domain. $Csd2_{121–308}$ was used for structure determination. (B) SEC-MALS data for two Csd2 protein samples. The black solid lines represent the measured molecular masses. The average molecular masses from MALS analyses are compared with the calculated masses in the table below the chromatography profiles.

Therefore, equilibrium sedimentation experiments were additionally carried out. The data measured at two different rotor speeds and at two different protein concentrations indicate that both $Csd2_{121–308}$ and $Csd2_{140–251}$ exist as reversible monomer-dimer (1x-2x) equilibrium in solution. The representative results, as shown in Figure 2-8, gave dissociation constants (K_d) of 4.9 μM for $Csd2_{121–308}$ and 45 μM for $Csd2_{140–251}$. Deletion of the helical domain was not sufficient for complete disruption of $Csd2_{140–251}$ dimerization but resulted in facilitated dissociation of $Csd2_{140–251}$ dimers, with about a nine-fold increase in the K_d value. This finding is in agreement with the crystal structure of $Csd2_{121–308}$ homodimer. On the basis of these results, one may expect that Csd2 can possibly interact with its close homolog such as Csd1 in a similar manner. I confirmed by SEC-MALS that $Csd2_{121–308}$ and $Csd1_{54–312}$ form a stable heterodimer, as described in detail below.

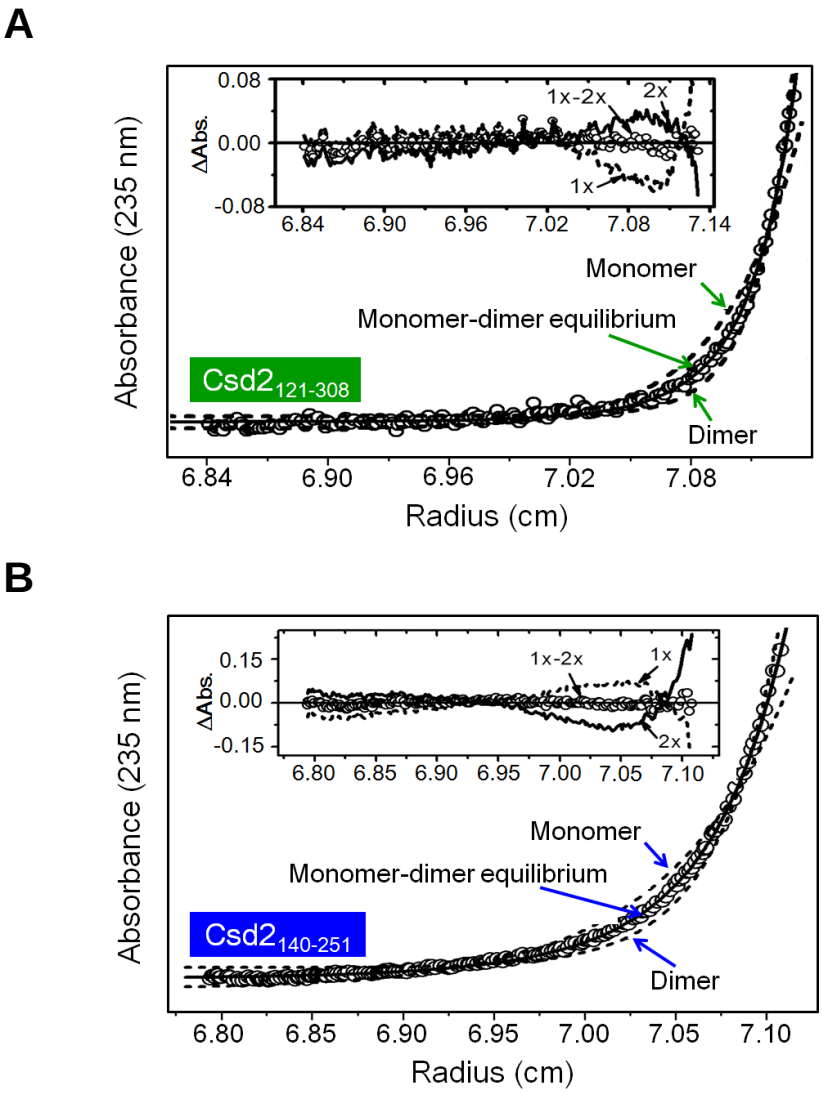


Figure 2-8. Equilibrium sedimentation to determine the oligomeric state of *H. pylori* Csd2 in solution

Equilibrium sedimentation data for two Csd2 protein samples. For $Csd2_{121–308}$ (A), the circles are experimental data measured at a speed of 35,000 rpm and 5.1 μM protein monomer concentration and the solid line is a fitting line for a reversible monomer-dimer equilibrium model. The two dotted lines are fitting lines for ideal homogeneous monomer and dimer models. Distributions of the residuals for monomer (dotted line), dimer (solid line), and reversible monomer-dimer equilibrium (circles) models are shown in the inset panel. For $Csd2_{140–251}$ (B), the circles are experimental data measured at a speed of 35,000 rpm and 14.5 μM protein monomer concentration and the solid line is a fitting line for a reversible monomer-dimer equilibrium model. The two dotted lines are fitting lines for ideal homogeneous monomer and dimer models. Distributions of the residuals for monomer (dotted line), dimer (solid line), and reversible monomer-dimer equilibrium (circles) models are shown in the inset panel. Sedimentation equilibrium data indicate that both $Csd2_{121–308}$ and $Csd2_{140–251}$ are in reversible monomer-dimer equilibrium in solution.

2.3.3. Csd1 and Csd2 can form a stable heterodimer in solution

To examine the possible complex formation between Csd1 and Csd2 by affinity chromatography, I overproduced four Csd1 constructs fused with one or two hexahistidine-containing tags and $Csd2_{121–308}$ without a fusion tag. Four possible complexes between Csd1 and Csd2 are purified by mixing the cell pellets that individually express either a Csd1 construct or the $Csd2_{121–308}$ construct, as described in Material and methods. The tested Csd1 constructs are (i) $Csd1_{54–312}$ fused with a C-terminal tag, (ii) $Csd1_{75–312}$ with a C-terminal tag, (iii) $Csd1_{91–312}$ with a C-terminal tag, and (iv) $Csd1_{125–312}$ with tags at both N- and C-termini. The complex formation between each of the above four Csd1 constructs and $Csd2_{121–308}$ could be readily identified by SDS-PAGE analysis (Figure 2-9). This result indicates that stable complexes of different Csd1 constructs and $Csd2_{121–308}$ can be formed and purified, when individually-expressed Csd1 and Csd2 proteins are present in roughly equal amounts. This implies that these Csd1-Csd2 complexes are more stable than the $Csd2_{121–308}$ homodimer.

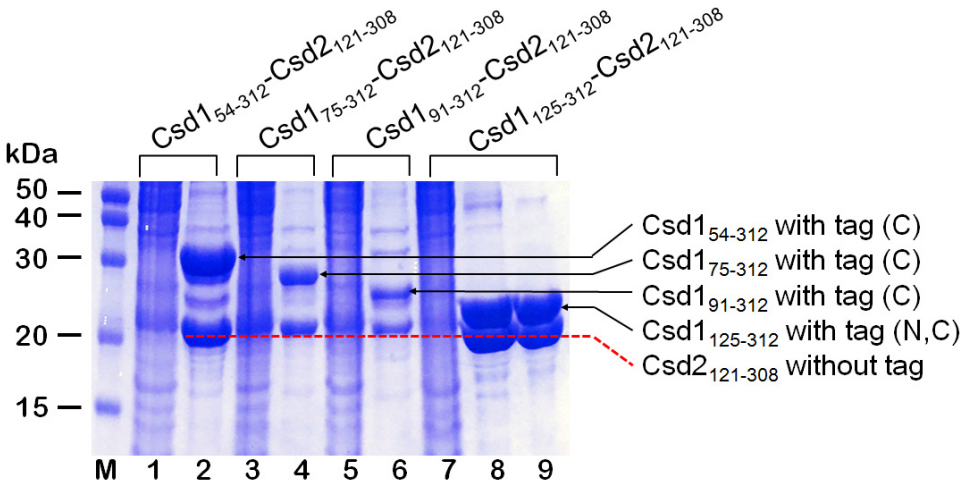
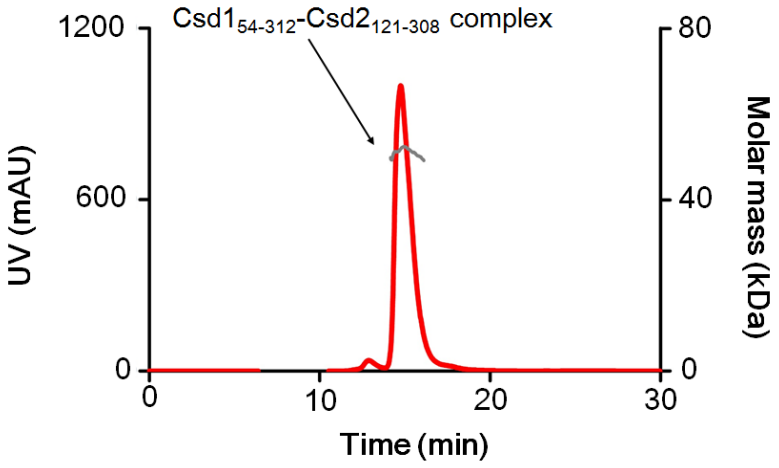


Figure 2-9. SDS-PAGE analysis of the Csd1 and Csd2 complex formation

Lane M: pre-stained protein ladder. Lanes 1 and 2 for $Csd1_{54\text{-}312}$-$Csd2_{121\text{-}308}$ complex: pooled fractions either unbound (lane 1) or bound (lane 2) to the affinity chromatography. Lane 3 (unbound) and lane 4 (bound) for $Csd1_{75\text{-}312}$-$Csd2_{121\text{-}308}$ complex. Lane 5 (unbound) and lane 6 (bound) for $Csd1_{91\text{-}312}$-$Csd2_{121\text{-}308}$ complex. Lane 7 (unbound) and lane 8 (bound) for $Csd1_{125\text{-}312}$-$Csd2_{121\text{-}308}$ complex. Lane 9: the peak fraction of the $Csd1_{125\text{-}312}$-$Csd2_{121\text{-}308}$ complex after size exclusion chromatography. This complex was crystallized for structure determination.

To establish the stoichiometry of the Csd1-Csd2 complexes, I estimated the molecular mass of the purified complex between $Csd1_{54–312}$ and $Csd2_{121–308}$ by SEC-MALS (Figure 2-10). The measured molecular mass of 51.5 kDa for the complex agrees well with the calculated total mass of $Csd1_{54–312}$ (30.8 kDa) and $Csd2_{121–308}$ (21.3 kDa). This result establishes that $Csd1_{54–312}$ and $Csd2_{121–308}$ bind in a 1:1 molar ratio. Additionally, I have determined the crystal structure of the 1:1 complex between $Csd1_{125–312}$ and $Csd2_{121–308}$, as described in more detail below.



Calculated MW (kDa)	MALS MW (kDa)
30.8 (Csd1$_{54-312}$) 21.3 (Csd2$_{121-308}$)	51.5 (±0.28 %)

Figure 2-10. SEC-MALS data for the Csd1$_{54–312}$-Csd2$_{121–308}$ complex

The red line represents the size exclusion chromatography profile. The grey line represents the measured molecular mass, whose average value agrees well with the calculated molecular mass of a 1:1 complex, as shown in the table below the chromatography profile.

2.3.4. Crystal structure of the heterodimer between $Csd1_{125–312}$ and $Csd2_{121–308}$

To reveal detailed interactions between Csd1 and Csd2, I purified four complexes ($Csd1_{54–312}$-$Csd2_{121–308}$, $Csd1_{75–312}$-$Csd2_{121–308}$, $Csd1_{91–312}$-$Csd2_{121–308}$, and $Csd1_{125–312}$-$Csd2_{121–308}$) and tried to crystallize them. Only the $Csd1_{125–312}$-$Csd2_{121–308}$ complex was crystallized in the monoclinic $P2_1$ space group (Table 2-3). I have solved the structure of this complex using two different sets of data collected from two different crystals at 2.27 and 2.35 Å resolutions, respectively (Csd1-Csd2 dimer I and Csd1-Csd2 dimer II in Table 2-4). In both heterodimer models, two copies of the heterodimer (AB dimer and CD dimer) are present in the asymmetric unit, with each heterodimer consisting of $Csd1_{125–312}$ (chain A or C) and $Csd2_{121–308}$ (chain B or D) in a 1:1 molar ratio (Figure 2-11). The two heterodimers within the asymmetric unit are related to each other by non-crystallographic two-fold symmetry, with Cα r.m.s. deviations of 1.86 Å and 2.15 Å for 332 and 334 residues for heterodimer models I and II, respectively. $Csd1_{125–312}$ and $Csd2_{121–308}$ chains within each of the heterodimers are related by pseudo two-fold symmetry due to their overall structural similarity (Cα r.m.s. deviations of 1.95–4.47 Å for 159–165 residues). The $Csd1_{125–312}$ and

$Csd2_{121–308}$ structures in dimer model I account for the following residues (Table 2-8): (i) 129–153 and 165–299 of chain A, (ii) 127–156 and 168–299 of chain C, (iii) 123–304 of chain B, and (iv) 123–297 of chain D. The $Csd1_{125–312}$ and $Csd2_{121–308}$ structures in dimer model II account for the following residues: (i) 128–154 and 161–299 of chain A, (ii) 128–154 and 169–300 of chain C, (iii) 122–297 of chain B, and (iv) 122–298 of chain D. The C-terminal residues 299–304 of $Csd2_{121–308}$ (chain B) in dimer model I are not involved in hetero-dimerization but are instead inserted into the substrate binding groove of $Csd1_{125–312}$ (chain C'), as described in more detail below.

A strong electron density was observed at the metal-binding site in each of $Csd1_{125–312}$ chains A and C in both heterodimer models I and II (Figure 2-11). As the key residues for the catalytic activity are well conserved in the Csd1 LytM domain, I have assigned the metal ion as Zn^{2+}. This assignment was confirmed by the anomalous difference electron density maps calculated using the anomalous data collected to 2.90 Å at the Zn^{2+} X-ray absorption edge of 1.2826 Å from a third $P2_1$ crystal (Table 2-6). In contrast, no electron density is observed at the corresponding site of $Csd2_{121–308}$ chains B and D in both heterodimer models I and II (Figure 2-11), indicating that Zn^{2+} is not bound.

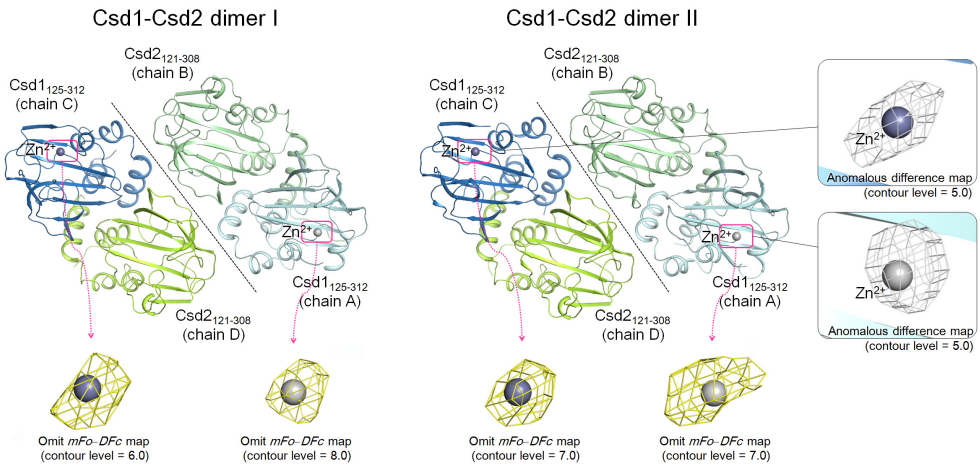


Figure 2-11. Two structures of $Csd1_{125–312}$-$Csd2_{121–308}$ heterodimer

Two structures of $Csd1_{125–312}$-$Csd2_{121–308}$ heterodimer (Csd1-Csd2 dimer I and dimer II) as determined using two different data sets are shown in ribbon diagram. Both heterodimer structures contain two copies of the heterodimer in the asymmetric unit: an AB dimer, formed by $Csd1_{125–312}$ (chain A, light blue) and $Csd2_{121–308}$ (chain B, light green), and a CD dimer, formed by $Csd1_{125–312}$ (chain C, colored in sky blue) and $Csd2_{121–308}$ (chain D, yellow-green). A black dotted line divides two copies of the heterodimer in the asymmetric unit, which are related by non-crystallographic two-fold symmetry. Anomalous difference electron densities for Zn^{2+} ions in chains A and C of dimer II are shown in white-grey mesh (right). The omit mF_o-DF_c map for Zn^{2+} ions in all four chains of Csd1 are shown in yellow mesh (bottom). No metal ion is bound to Csd2 chains.

Table 2-8. Modeled residues of Csd2 homodimer and Csd1-Csd2 heterodimer structures

Structure	*Chain A* ($Csd1_{125–312}$)	*Chain B* ($Csd2_{121–308}$)	*Chain C* ($Csd1_{125–312}$)	*Chain D* ($Csd2_{121–308}$)
Csd1-Csd2 dimer I	**Gly(129)–Ala(153)** **Val(165)–Gln(299)**	**Asn(123)–Ala(304)**	**Ile(127)–Asn(156)** **Asn(168)–Gln(299)**	**Asn(123)–Gln(297)**
Csd1-Csd2 dimer II	**Thr(128)–Ala(154)** **Pro(161)–Gln(299)**	**Asp(122)–Gln(297)**	**Thr(128)–Ala(154)** **His(169)–Lys(300)**	**Asp(122)–Glu(298)**

Structure	*Chain A* ($Csd2_{121–308}$)
Csd2-Csd2 dimer	H–**Leu(121)–Asp(301)**

The amino acid sequences of $Csd1_{125–312}$ and $Csd2_{121–308}$ are denoted by bold-faced three-letter codes and the fusion tag in Csd2-Csd2 dimer structure are denoted by single letter code.

The $Csd1_{125–312}$ monomer (chain C in heterodimer model I) is composed of four α-helices, one 3_{10}-helix, and ten β-strands in the order α1-β1-β2-β3-β4-η1-β5-β6-β7-β8-β9-β10-α2-α3-α4 (Figure 2-12). Similarly to $Csd2_{121–308}$, the structure of $Csd1_{125–312}$ can be divided into two domains: a helical domain (residues Ile127–Asp142 and Asn263–Gln299) consisting of four helices (α1–α4) and the C-terminal LytM domain (residues Tyr143–Ile262) consisting of ten strands (β1–β10) (Figure 2-12). As in $Csd2_{121–308}$, the C-terminal LytM domain is inserted between helices α1 and α2 of the helical domain. In the LytM domain of Csd1 (chain C in heterodimer model I, with the canonical Zn^{2+}-coordination), a central anti-parallel β-sheet (β1↑-β2↓-β9↑-β6↓-β5↑-β4↓-β7↑) forms a two-layered sandwich with a smaller anti-parallel β-sheet (β3↓-β8↑-β6↓) (Figure 2-12). The long, highly curved strand β6 is shared between these two sheets. A short strand β10 forms a mini, anti-parallel β-sheet (β9↑-β10↓) with the C-terminal side of β9 of the central β-sheet. As expected, the Csd1 LytM domain exhibits a high level of structural similarity with other LytM domains of the M23 peptidase family, with Z-scores up to 18.5 (Table 2-9).

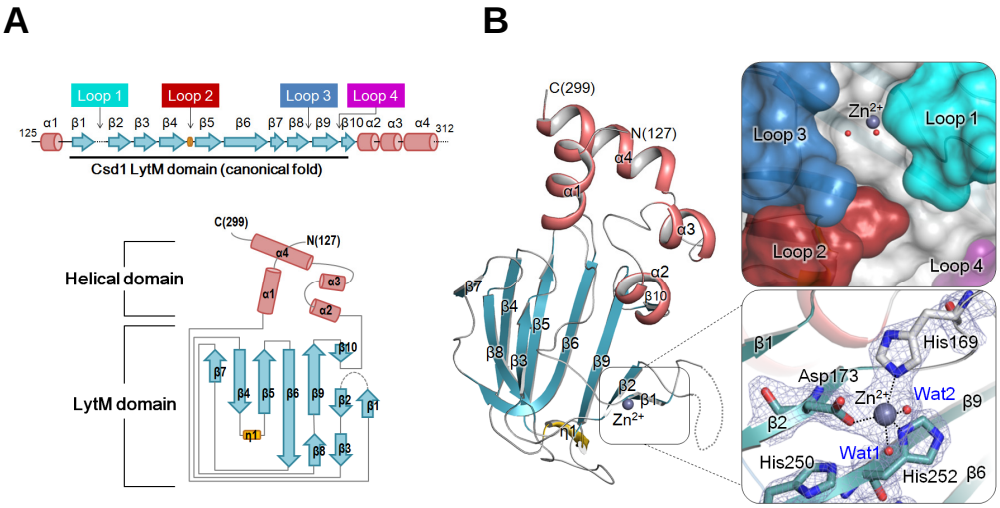


Figure 2-12. Overall monomer structure of *H. pylori* Csd1$_{125–312}$

(A) Schematic representation of secondary structures of $Csd1_{125–312}$ and topology diagram of $Csd1_{125–312}$. Secondary structures have been defined by the STRIDE program (Heinig and Frishman, 2004). α-Helices, β-strands, 3_{10}-helices, and loops are shown as cylinders (colored in light pink), arrows (blue-green), flat cylinders (yellow), and solid lines (grey), respectively. Loop 1 (β1-β2 loop; cyan), Loop 2 (β4-β5 loop; red), Loop 3 (β8-β9 loop; skyblue), and Loop 4 (β9-β10 loop; purple) form the substrate-binding groove of the Csd1 LytM domain. Dotted lines indicate disordered regions. (B) Ribbon diagram of $Csd1_{125–312}$ monomer structure (chain C of Csd1-Csd2 dimer I), colored as in the topology diagram in (A). Close-up views on the right represent the surface representation of the substrate-binding groove formed by four loops of the LytM domain (top) and canonical Zn^{2+}-coordination with three protein ligands and two water molecules (bottom). Dark grey and red spheres represent a Zn^{2+} ion and water molecules, respectively. Side chains of the Zn^{2+}-coordinating residues (His169, Asp173, and His252) are shown in stick models. Black dotted lines denote penta-coordination of the Zn^{2+} ion. The electron density for the Zn^{2+}-bound active site in $2mF_o - DF_c$ map (grey colored mesh) are shown at the 1.0 σ level.

Table 2-9. Structural similarity searches with the Csd1 LytM domain

	Protein name		PDB code	R.m.s. deviation	Z-score	Sequence identity
Csd1 LytM domain	Three-domain protein	Lysostaphin peptidase (VC0503) from *Vibrio cholera*	2GU1	1.2 Å (106 Cα)	18.5	43%
		Outer-membrane protein (NMB0315) from *Neisseria meningitidis*	3SLU	1.3 Å (104 Cα)	18.2	46%
		Cell shape determinant 3 (Csd3; HP0506) from *Helicobacter pylori*	4RNY	1.0 Å (100 Cα)	17.5	36%
	Pro-protein	Glycylglycine endopeptidase (LtyM) from *Staphylococcus aureus*	2B13	1.5 Å (106 Cα)	17.7	32%
		Lysostaphin from *Staphylococcus simulans*	4QPB	1.5 Å (105 Cα)	17.2	39%
	Non-peptidase protein	Murein hydrolase activator (EnvC) from *Escherichia coli*	4BH5	1.2 Å (104 Cα)	17.8	29%
		Stage II sporulation protein Q (SpollQ) from *Bacillus subtilis*	3UZ0	1.5 Å (106 Cα)	15.6	34%

All four $Csd2_{121–308}$ structures (chains B and D) in the two heterodimer models are essentially identical, with Cα r.m.s. deviations of 0.29–0.42 Å for 175–176 residues. They also do not differ much from the $Csd2_{121–308}$ structure in the homodimer model, with Cα r.m.s. deviations of 0.44–1.05 Å for 175 –179 residues. Interestingly, however, the segments of $Csd1_{125–312}$ around the metal-binding site display diverse structures among different chains of $Csd1_{125–312}$ in the two heterodimer models. More specifically, four $Csd1_{125–312}$ structures (chains A and C) in the two heterodimer models adopt three different modes of metal coordination in the LytM domain (Figure 2-13A). Two $Csd1_{125–312}$ structures in dimer model I (chain C) and model II (chain C) are virtually identical with a Cα r.m.s. deviation of 0.44 Å for 158 residues and share the canonical metal coordination. Two $Csd1_{125–312}$ structures in dimer model I (chain A) and model II (chain A) deviate from chains C (Cα r.m.s. deviations of 2.60–3.04 Å for 156–159 residues) but they also differ from each other (Cα r.m.s. deviation of 2.04 Å for 160 residues). As a consequence, I observe two different types of non-canonical metal-coordination by the Csd1 LytM domain (Figure 2-13B), as further discussed below. The observed structural differences around the metal-binding sites of $Csd1_{125–312}$ have little effect on the heterodimerization pattern, because the

metal-binding site is well separated from the interface between $Csd1_{125–312}$ and $Csd2_{121–308}$ in the heterodimer.

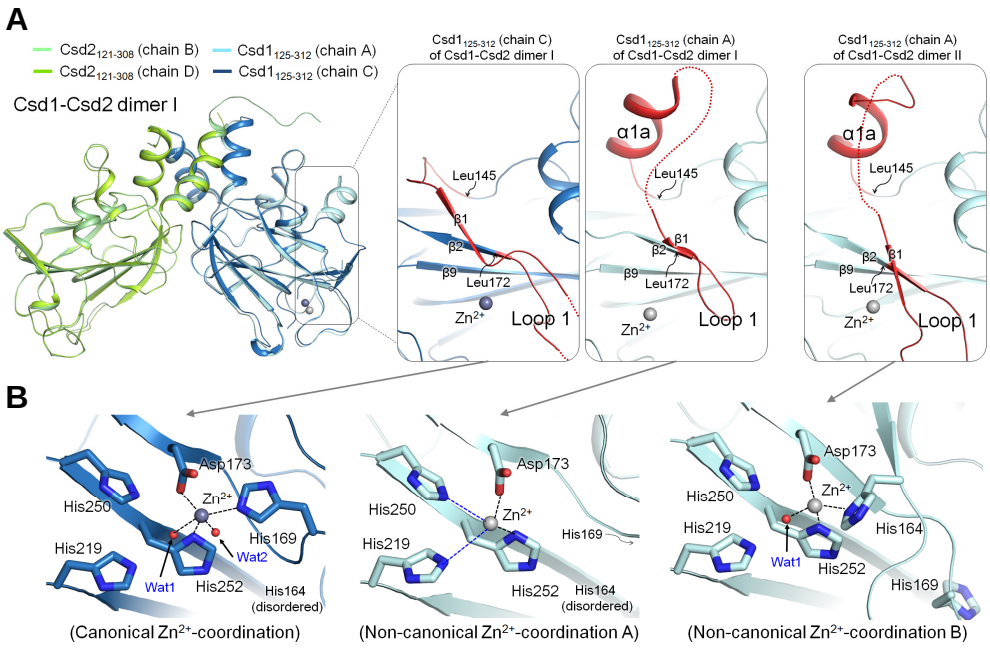


Figure 2-13. Three different types of Zn^{2+}-coordination are observed in the Csd1 LytM domains

(A) Two copies of the $Csd1_{125–312}$-$Csd2_{121–308}$ heterodimer (AB dimer and CD dimer) in the structure of Csd1-Csd2 dimer I are superimposed and shown in ribbon diagram (left). Detailed views of the region covering the sequence of Leu145–Leu172 (colored in red) in (i) $Csd1_{125–312}$ chain C of the Csd1-Csd2 dimer I (left), (ii) $Csd1_{125–312}$ chain A of the Csd1-Csd2 dimer I (middle), and (iii) $Csd1_{125–312}$ chain A in the Csd1-Csd2 dimer II (right) are shown in the black boxes. The region covering Leu145–Leu172 is structurally most divergent among different chains of Csd1. The dotted lines represent disordered loops. An extra α-helix (labeled as α1a) is formed in $Csd1_{125–312}$ chain A in both structures of Csd1-Csd2 dimer I and dimer II. Four chains of $Csd1_{125–312}$ have different lengths of Loop 1. Loop 1 in $Csd1_{125–312}$ chain A of the Csd1-Csd2 dimer I (middle) is much shorter than those in chain C of the Csd1-Csd2 dimer I structure (left) and chain A of the Csd1-Csd2 dimer II structure (right).

(B) Detailed views of three different types of Zn^{2+}-coordination by the LytM domain of Csd1. The central β-sheet is shown in ribbon diagram, with Zn^{2+}-coordinating residues (His164, His169 and Asp173 of the HxxxD motif, His250 and His252 of the HxH motif, and highly conserved His219) in stick models. Dotted lines represent direct Zn^{2+}-coordination or close contacts. In the canonical coordination, the Zn^{2+} ion is coordinated by three conserved

ligands (His169, Asp173, and His252) and two water molecules (Wat1 and Wat2). His164 is disordered in this model. In non-canonical coordination A, the Zn^{2+} ion is coordinated by Asp173, His219, His250, and His252. His169 is far away from the Zn^{2+} ion and is not included in this figure. Again, His164 is disordered in this model. In non-canonical coordination B, the Zn^{2+} ion is coordinated by His164 and two conserved ligands (Asp173 and His252) and a water molecule (Wat1). His169 is moved away from the Zn^{2+} ion. His164 replaces His169 in canonical coordination to act as the metal ligand.

In the crystal structure of the heterodimer between $Csd1_{125–312}$ and $Csd2_{121–308}$, the pair of helices α1 and α4 from the helical domain of $Csd1_{125–312}$ pack against the corresponding pair of helices α1 and α3 from the helical domain of $Csd2_{121–308}$ to form a four helix bundle, in which one $Csd1_{125–312}$ monomer essentially replaces a $Csd2_{121–308}$ monomer within the $Csd2_{121–308}$ homodimer (Figure 2-14). Buried surface areas per monomer of $Csd1_{125–312}$ and $Csd2_{121–308}$, as calculated using the PISA sever (Krissinel and Henrick, 2007), differ by only 1.1–4.4%. Four heterodimers have buried surface areas per monomer (averaged over $Csd1_{125–312}$ and $Csd2_{121–308}$) of 1,310, 1,290, 1,400, and 1,440 $Å^2$ for AB and CD heterodimer models I and II, respectively. These values are slightly smaller than that for the $Csd2_{121–308}$ homodimer, because more N- and C-terminal residues of $Csd2_{121–308}$ as well as $Csd1_{125–312}$ are disordered and invisible in the heterodimers between $Csd1_{125–312}$ and $Csd2_{121–308}$ than in the $Csd2_{121–308}$ homodimer (Figure 2-14 and Table 2-8). An extended hydrophobic region is found around the center of the dimer interface in both the $Csd2_{121–308}$ homodimer and the heterodimer between $Csd1_{125–312}$ and $Csd2_{121–308}$ (Figure 2-14). However, surface charge distributions of the surrounding regions in $Csd1_{125–312}$ and $Csd2_{121–308}$ are strikingly different. Highly negatively-charged surfaces surround the hydrophobic interface of the $Csd2_{121–308}$ homodimer, whereas largely

positively-charged surfaces surround the hydrophobic interface of $Csd1_{125–312}$ of $Csd1_{125–312}$ and $Csd2_{121–308}$ heterodimer (Figure 2-14). The interface between $Csd1_{125–312}$ and $Csd2_{121–308}$ harbors a network of hydrogen bonds and salt bridges (Figure 2-15A). Asp292 on helix α3 of $Csd2_{121–308}$ forms three salt bridges with Arg137 on helix α1 of $Csd1_{125–312}$ and His130 on helix α1 of $Csd2_{121–308}$ interacts with Gln299 on helix α4 of $Csd1_{125–312}$ (Figure 2-15B). A hydrogen bond exists between the side chains of Gln288 of $Csd2_{121–308}$ and Arg296 of $Csd1_{125–312}$. The main chain of Lys224 of $Csd2_{121–308}$ is hydrogen bonded to Arg286 of $Csd1_{125–312}$. These hydrogen bonds are formed between strand β7 of Csd2 LytM domain and helix α4 of Csd1 helical domain (Figure 2-15C). Two arginine residues (Arg286 and Arg296) on helix α4 of $Csd1_{125–312}$ correspond to negatively charged residues Glu282 and Asp292 of $Csd2_{121–308}$, respectively (Figures 2-15D and 2-15E).

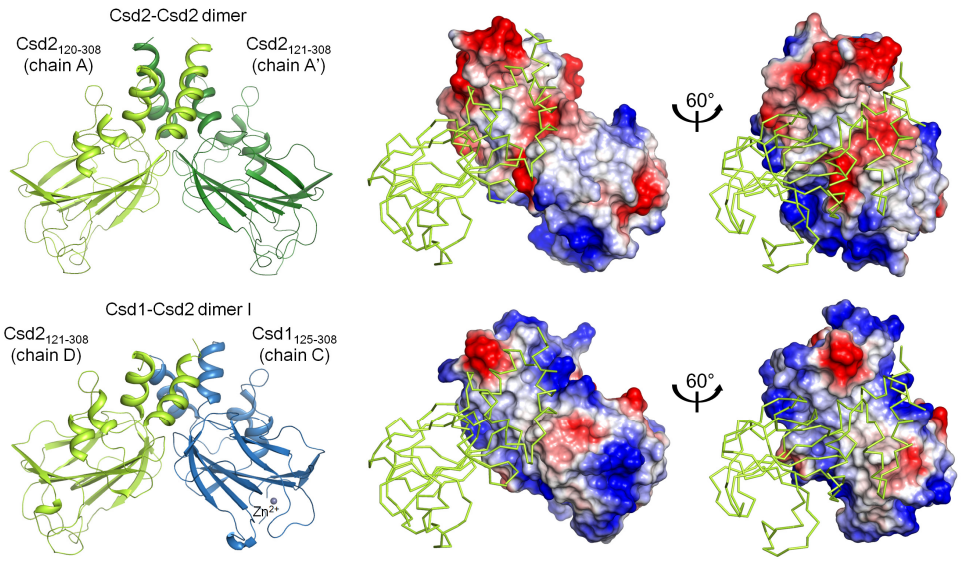


Figure 2-14. Comparison of $Csd2_{121–308}$ homodimer and $Csd1_{125–312}$-$Csd2_{121–308}$ complex

$Csd2_{121–308}$ homodimer (top left) and the $Csd1_{125–312}$-$Csd2_{121–308}$ complex (CD dimer in heterodimer model I) (bottom left) are shown in ribbon diagrams; they are colored as in Figure 2-6A and Figure 2-11, respectively. Electrostatic surface diagrams of $Csd2_{121–308}$ chain A' of Csd2 homodimer and $Csd1_{125–312}$ chain C of the $Csd1_{125–312}$-$Csd2_{121–308}$ heterodimer are shown on the right. Highly negatively-charged surfaces surround the hydrophobic interface of the $Csd2_{121–308}$ homodimer, whereas largely positively-charged surfaces surround the hydrophobic interface of $Csd1_{125–312}$ of the $Csd1_{125–312}$-$Csd2_{121–308}$ heterodimer.

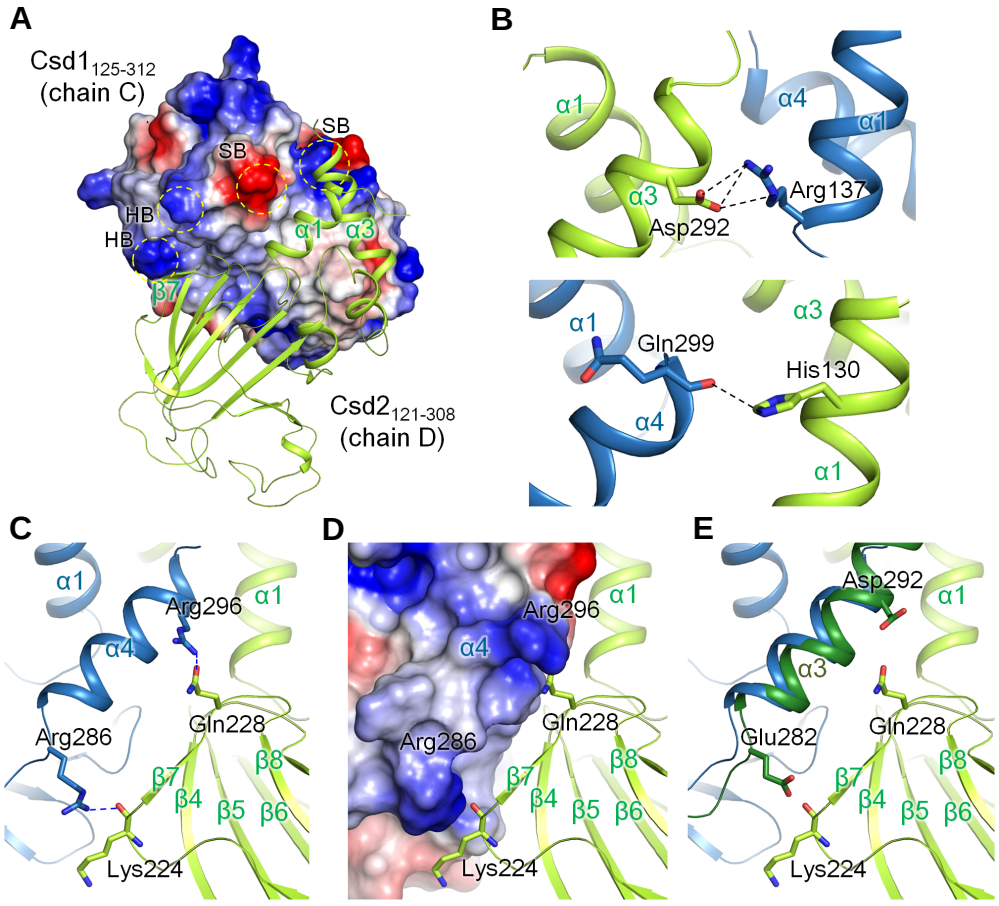


Figure 2-15. Dimer interface within the Csd1$_{125–312}$-Csd2$_{121–308}$ complex

(A) The heterodimer interface between $Csd1_{125–312}$ (chain C, shown in the electrostatic surface diagram) and $Csd2_{121–308}$ (chain D, shown in ribbon diagram). The AB dimer of heterodimer I, and AB and CD dimers of heterodimer II have highly similar interfaces. Hydrogen bonding (HB) and salt bridge (SB) interactions are indicated by the dotted yellow circles.

(B and C) Detailed views of the salt bridges (B) and hydrogen bonds (C) are indicated by dotted lines. $Csd1_{125–312}$ (chain C) is in sky blue, while $Csd2_{121–308}$ (chain D) is in yellow-green, as in Figure 2-14.

(D) The electrostatic surface diagram represents the positively charged surface of α4 helix in $Csd1_{125–312}$ (chain C) in the hetero-dimer interface. $Csd2_{121–308}$ (chain D) is shown in ribbon diagram (yellow-green).

(E) Superimposition of α3 helix in $Csd2_{121–308}$ (darker green) onto α4 helix of $Csd1_{125–312}$ (yellow-green). Glu282 and Asp292 on the $Csd2_{121–308}$ α3 helix, which structurally correspond to Arg286 and Arg296 of $Csd1_{125–312}$, respectively, are shown in stick models.

2.3.5. LytM domains of Csd1 and Csd2 and comparisons with Csd3 LytM domain

Amino acid sequences of Csd1, Csd2, and Csd3 from *H. pylori* 26695 strain are well aligned over their LytM domains, whose overall structures are also similar to each other (Figures 2-16 and 2-17). However, outside their LytM domains, Csd1 and Csd2 are not well aligned with Csd3 in the amino acid sequences as well as in their structures (An et al., 2015) (Figure 2-16). Two characteristic motifs (HxxxD and HxH) of the LytM domain are conserved in Csd1 and Csd3, whereas the first histidines of these motifs in Csd2 are substituted with Glu165 and Lys246, respectively (Figure 2-16).

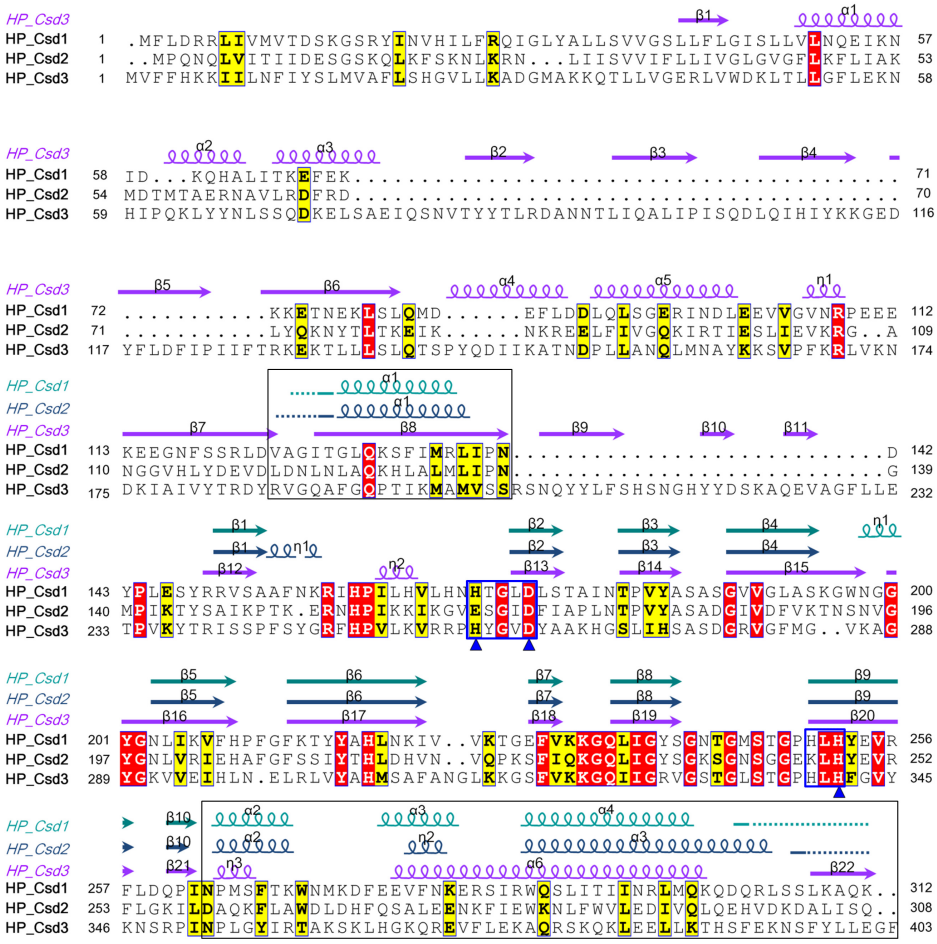


Figure 2-16. Sequence alignment of cell shape-determinant proteins Csd1, Csd2, and Csd3 from *H. pylori* 26695 strain

Sequences of Csd1 (HP_Csd1; SWISS-PROT accession code: O26068), Csd2 (HP_Csd2; O26069), and Csd3 (HP_Csd3; O25247) were aligned using Clustal Omega (Sievers et al., 2011) and the alignment figure was drawn using ESPript (Gouet et al., 2003; http://espript.ibcp.fr). The secondary structures are presented above the aligned sequences. Two large black boxes indicate the helical domains in Csd1 and Csd2. In Csd3, they correspond to part of Domain 2 and the C-terminal helical region that associates with Domain 2 (An et al., 2015). LytM domains are found between these black boxes. Two small blue boxes indicate the conserved HxxxD and HxH motifs in the LytM domains, with blue triangles corresponding to the Zn^{2+}-coordinating residues of Csd1 and Csd3.

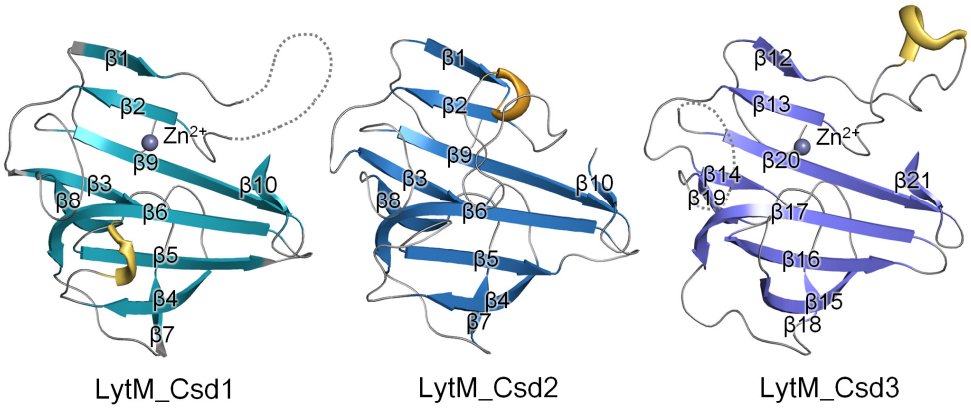


Figure 2-17. Structural comparison of three LytM domains

Three LytM domains from *H. pylori* Csd1 (chain C in heterodimer model I), Csd2 (in homodimer), and Csd3 (in Form 1 crystals) (An et al., 2015) are shown as ribbon diagram. Dotted grey lines indicate disordered loops in Csd1 and Csd3 LytM domains. 3_{10}-Helices are colored in yellow. A Zn^{2+} ion is bound to the LytM domains of Csd1 and Csd3, whereas no Zn^{2+} ion is bound to the Csd2 LytM domain.

Therefore, it is expected that Csd2 possesses a degenerate active site in its LytM domain, like other non-peptidase members of the M23B metallopeptidase family. In agreement with this expectation, no metal ion is bound to the active site in the Csd2 LytM domain in structures of both Csd2 homodimer and Csd1-Csd2 heterodimer (Figures 2-5B, 2-11, and 2-17). Despite the absence of a metal ion in the active site, the Csd2 LytM domain possesses a substrate-binding groove formed by four loops (Loop 1; β1-β2, Loop 2; β4-β5, Loop 3; β8-β9, and Loop 4; β9-β10) on the central β-sheet (Figures 2-5 and 2-17). Loop 1 is folded over the degenerate active site and covers it. Similar substrate-binding sites are present in other M23 family member proteins (Spencer et al., 2010).

The four $Csd1_{125–312}$ monomers in two heterodimer models I and II show three different modes of metal-coordination in the LytM domain: the canonical metal-coordination (for both chains C in heterodimer models I and II) and two other non-canonical modes of metal-coordination (i.e., non-canonical coordination A for chain A of heterodimer model I and non-canonical coordination B for chain A of heterodimer model II) (Figure 2-13). The observed diversity in metal-coordination by the LytM domains of four $Csd1_{125–312}$ monomers in two heterodimer models I and II, likely resulting from differences in their environment within the crystals, suggests that the

segment of $Csd1_{125–312}$ covering residues Leu145–Leu172 possesses high flexibility and structural plasticity in forming the metal-binding site. This segment includes strand β1 and Loop 1 (Figure 2-13A). In the structure of Csd3 LytM domain (An et al., 2015), the floor of the substrate-binding groove is built upon the central β-sheet and the walls of the active site are made up of four loops: Loop I (β12-β13 loop), Loop II (β15-β16 loop), Loop III (β19-β20 loop), and Loop IV (β20-β21 loop) (Figure 2-17). In Form 1 crystals of Csd3, Loop I is involved in crystal packing interactions by forming sulfate-mediated salt bridges. In Form 2 crystals, it is not involved in the crystal packing and is disordered. Therefore, it was concluded that the observed structural variation of Loop I is largely due to the difference in crystal packing and also due to its inherent flexibility (An et al., 2015). Despite the flexibility of Loop I in Csd3, only the canonical Zn-coordination was observed.

In the LytM domain of $Csd1_{125–312}$ (chain C of heterodimer model I), with the canonical Zn^{2+}-coordination, the central β-sheet anchors the catalytic residues, which are grouped around the Zn^{2+} ion. The zinc-bound active site is located in a substrate-binding groove that is made up of four loops (Loop 1; β1-β2, Loop 2; β4-β5, Loop 3; β8-β9, and Loop 4; β9-β10) (Figures 2-12 and 2-17). Loop 1 is partially disordered and Loop 2 contains a

3_{10}-helix (η1). The Zn^{2+} ion is penta-coordinated in slightly distorted trigonal bipyramidal geometry with the expected three ligands (His169 and Asp173 of the **H**xxx**D** motif and His252 of the Hx**H** motif) and two water molecules (Figures 2-12B and 2-13A). The metal-ligand atom distances are in the range of 2.0–2.5 Å, consistent with typical Zn^{2+} ion-ligand atom distances. The two water molecules are assigned as Wat1 (Wat1–Zn^{2+} 2.5 Å), which is coordinated by conserved histidine residues (His 250 of the **H**xH motif and His219), and Wat2 (Wat2–Zn^{2+} 2.3 Å) (Figures 2-12B and 2-13A). The observed penta-coordination is coincident with a catalytic mechanism proposed for LasA from *P. aeruginosa* (Spencer et al., 2010) and other LytM domains (Sabala et al., 2014). In the tartrate-uncomplexed structure of LasA from *P. aeruginosa*, the Zn^{2+} ion is also penta-coordinated by three conserved metal ligands and two water molecules. The two water molecules are assigned as Wat1 (Wat1–Zn^{2+} 2.1 Å), which is oriented by interactions with two conserved histidine residues (His120 of the **H**xH motif and His80), and Wat2 (Wat2–Zn^{2+} 2.7 Å) (Spencer et al., 2010). In the proposed mechanism, the substrate carbonyl oxygen displaces Wat2, while Wat1 functions as a nucleophile to attack the polarized carbonyl bond (Spencer et al., 2010; Sabala et al., 2014). In the inactivated structure of Csd3 (An et al., 2015), the Zn^{2+} ion in the active site of the LytM domain is tetra-coordinated,

with His259 and Asp263 of the **H**xxx**D** motif, His341 of the Hx**H** motif, and Glu74 from the α3 helix of Domain 1. Two water molecules necessary for the peptidase activity are replaced by side chain oxygens from Glu74: Wat1 by Glu74$^{\varepsilon 2}$ (2.0 Å from the Zn^{2+} ion) and Wat2 by Glu74$^{\varepsilon 1}$ (2.9 Å from the Zn^{2+} ion) (An et al., 2015).

Two other non-canonical modes of Zn^{2+}-coordination in the $Csd1_{125–312}$ structures of heterodimer models I and II have not been reported previously. In both non-canonical coordination A and B, one of the three metal ligands (His169 of the HxxxD motif) does not participate in the Zn^{2+}-coordination and an additional α-helix (α1a) is formed in the linker between α-helix α1 of the helical domain and β-strand β1 of the LytM domain (Figure 2-13A), making Loop 1 shorter than in the canonical coordination. However, residues between helix α1a and Loop 1 are disordered in the crystal. In the non-canonical coordination A, the Zn^{2+} ion has a high B-factor of 82.4 $Å^2$, possibly due to partial occupancy. Thus, the Zn^{2+}-ligand atom distances are longer than usual (2.6 Å and 3.0 Å to Asp173 of the Hxxx**D** motif, 2.9 Å to His250 of the **H**xH motif, and 3.3 Å to His252 of the Hx**H** motif, and 3.3 Å to His219 (Figure 2-13B). In the non-canonical coordination B, the Zn^{2+} ion is tetrahedrally coordinated and has a lower B-factor of 43.8 $Å^2$, suggesting full occupancy. The Zn^{2+}-ligand atom distances are usual (1.9 Å to Asp173

of the Hxxx**D** motif, 2.2 Å to His252 of the Hx**H** motif, 2.4 Å to His164, and 2.1 Å to a water molecule (Wat1) (Figure 2-13B). In Csd1, the Zn^{2+}-coordinating His164 extends the HxxxD motif, resulting in the $H_{(164)}xxxxH_{(169)}xxxD_{(173)}$. His164 is not conserved in *H. pylori* Csd2 and Csd3. This may explain why non-canonical coordination is not observed for Csd2 and Csd3.

2.3.6. The C-terminal tail of Csd2 occupies the substrate-binding groove of Csd1 LytM domain in the crystal

Interestingly, during refinement of the heterodimer model I, we observed an extra electron density which extends from the C-terminal residue Glu298 of $Csd2_{121–308}$ (chain B') from an adjacent asymmetric unit of the crystal and runs along the groove between Loops 1 and 3 of active site of $Csd1_{125–312}$ (chain C with the canonical Zn^{2+}-coordination) (Figure 2-18A). We modeled this electron density as the C-terminal tail sequence HVDKDA of $Csd2_{121–308}$, encompassing residues 299–304 (Figure 2-18B). The electron density is weaker in the middle of the sequence HVDKDA. A similar electron density is absent in the active site of other $Csd1_{125–312}$ chains of heterodimer models I and II. This may be due to the fact that the HVDKDA

sequence differs from the sequence of the physiological substrate of Csd1. Nevertheless, the observed binding of the Csd2 tail sequence in the active site of Csd1 may define the substrate-binding site (P3-P2-P1-P1'-P2'-P3') of the Csd1 LytM domain, because the central peptide bond between Asp301 and Lys302 of the bound HVDKDA sequence is mimicking the peptide bond of the physiological substrate of Csd1. The interactions between the C-terminal tail residues of $Csd2_{121–308}$ and $Csd1_{125–312}$ are primarily mediated by the main chain atoms between Asp301 and Asp303 of $Csd2_{121–308}$, and Wat2 in the Zn^{2+}-coordination sphere of $Csd1_{125–312}$ (Figure 2-18C). The Asp301–Asp303 segment of the Csd2 tail covers the Csd1 active site; Wat2, one of the two Zn^{2+}-coordinating water molecules in the Csd1 active site, makes three interactions with the Csd2 tail [Asp301 carbonyl oxygen (2.99 Å), Lys302 amide nitrogen (with 2.89 Å), and Asp303 amide nitrogen (with 2.97 Å)] (Figure 2-18C). Both carbonyl oxygen and amide nitrogen atoms of Csd2 Val300 interact with Thr247 on Loop 3 of Csd1, while the carbonyl oxygen of Csd2 Asp303 interacts with Thr170 on Loop 1 of Csd1.

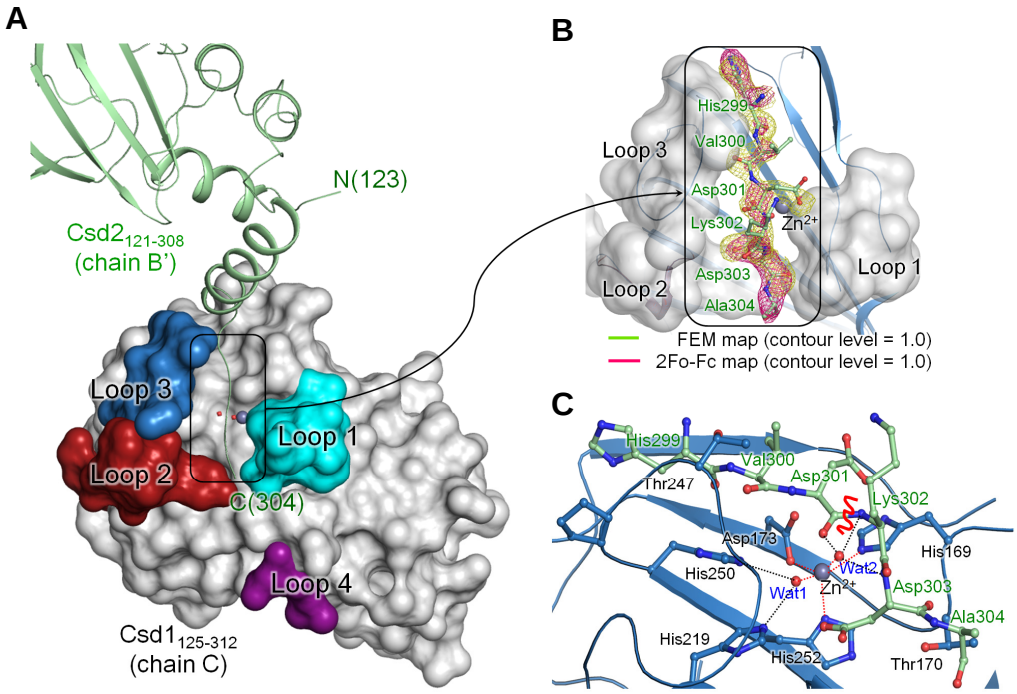


Figure 2-18. The C-terminal tail of Csd2 (chain B') is bound to the substrate-binding groove in the LytM domain of Csd1 (chain C in heterodimer model I)

(A) In the structure of Csd1-Csd2 dimer I, the C-terminal residues (His299–Ala304) of $Csd2_{121–308}$ from an adjacent asymmetric unit (chain B' shown in ribbon diagram) occupy the substrate-binding groove of the LytM domain in $Csd1_{125–312}$ (chain C shown in surface diagram). Four loops of Csd1 LytM domain that form the substrate-binding groove are labeled and colored as in Figure 2-12. The ribbon diagram is colored as in Figure 2-11.

(B) Close-up view of the Csd2 C-terminal tail residues located in the substrate-binding groove of Csd1 LytM domain. The Csd2 tail residues (enclosed in the black box) are shown in a stick model, with the electron density shown in mesh. The electron density for the Csd2 tail in the feature-enhance map (FEM) calculated by using PHENIX program (Afonine et al., 2015) (lime colored mesh) and $2mF_o - DF_c$ map (magenta colored mesh) are shown at the 1.0 σ level.

(C) A detailed view of the interactions between the Csd2 tail residues and the substrate-binding groove of the Csd1 LytM domain (shown in ribbon diagram, colored as in Figure 2-11). Both main chains and side chains of the Csd2 tail residues are shown in a stick model, with the candidate peptide bond that might be cleaved by the enzymatic activity of Csd1 is indicated by a red wavy line. Side chains of $Csd1_{125–312}$ residues interacting with the Csd2

tail residues are shown in a stick model. Grey and red spheres represent the Zn^{2+} ion and water molecules, respectively. Zn^{2+}-coordination (canonical) and hydrogen bonds with waters are indicated by red and black dotted lines, respectively.

The direction of the Csd2 peptide bound to Csd1 is consistent with the recent crystal structure of *Staphylococcus aureus* LytM in complex with tetraglycine phosphinate, a transition state analog of the penta-glycine bridge of peptidoglycan cross-links (Grabowska et al., 2015). This work helps to identify the substrate recognition and binding as well as describe details of the catalytic mechanism for cleaving the penta-glycine bridge by *S. aureus* LytM (Grabowska et al., 2015). The two Zn^{2+}-coordinating water molecules of $Csd1_{125–312}$ (chain C) superimpose nicely with two oxygen atoms of the phosphinate group in *S. aureus* LytM; they also superimpose well with two side chain oxygen atoms of Glu74 from the α3 helix of Domain 1 in *H. pylori* Csd3 (Figure 2-19). The oxyanion intermediate of the reaction catalyzed by *S. aureus* LytM was suggested to be stabilized by Tyr204 (Grabowska et al., 2015). Sequence alignment indicates that this tyrosine residue is replaced by a histidine residue (His160) in *H. pylori* Csd1, as well as in Csd2 and Csd3 (His156 in Csd2 and His250 in Csd3) (Figure 2-20). His160 of Csd1 is part of the disordered region in Loop 1 and is not included in the model. His156 of Csd2 is in proximity of the active site. However, His250 of Csd3 is located on Loop I, which protrudes from the LytM domain and is far from the active site.

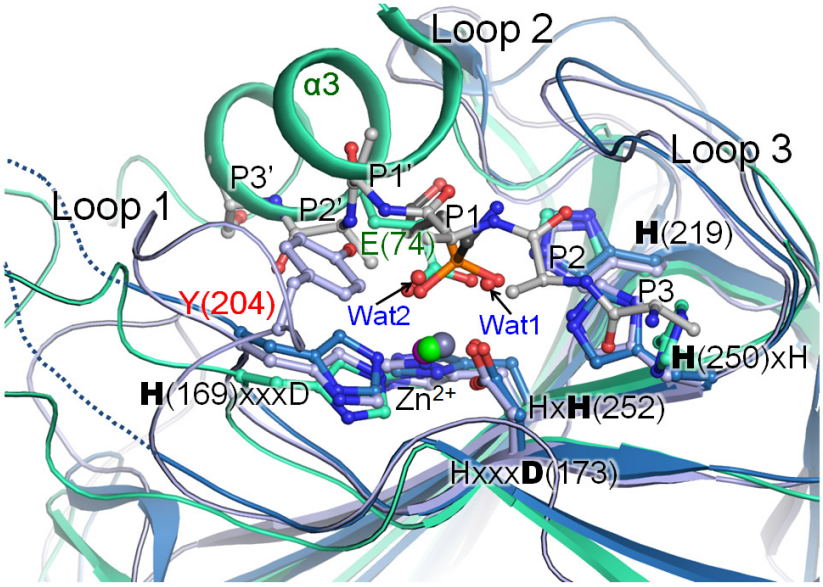


Figure 2-19. Superimposition of the substrate-binding groove in three LytM domains

Superimposition of LytM domains in *H. pylori* Csd1 (skyblue; chain C in heterodimer model I), *H. pylori* Csd3 (light green; PDB code, 4RNY), and *S. aureus* LytM bound with tetraglycine phosphinate (purple; PDB code, 4ZYB) shows that the two water molecules (Wat1 and Wat2) of $Csd1_{125–312}$ chain C of heterodimer I overlap nicely with side chain oxygen atoms of Glu74 (labeled in light green) from the helix α3 of the inhibitory domain 1 in Csd3 and also with those of the phosphinate molecule (black). The Csd2 tail is simplified as a poly-alanine model (grey). The bound Zn^{2+} ions are indicated by grey, purple, and green spheres for *H. pylori* Csd1, *H. pylori* Csd3, and *S. aureus* LytM, respectively. Two dotted lines represent the disordered regions in Loop 1 of Csd1. The metal-coordinating residues in the **H(169)**xxx**D(173)** and Hx**H(252)** motifs and the conserved catalytic residues in the **H(250)**xH motif and an additional catalytic histidine residue **H(219)** of Csd1, as well as corresponding residues of *H. pylori* Csd3 and *S. aureus* LytM, are shown in a stick model. Tyr204 (labeled in red) of *S. aureus* LytM is shown in a stick model.

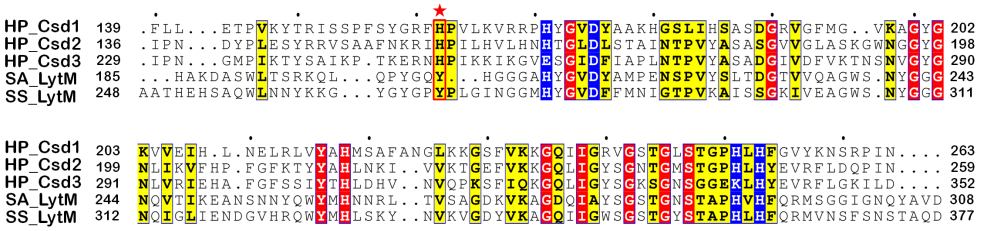


Figure 2-20. Sequence alignment of LytM domains

Sequence alignment of LytM domains in Csd1, Csd2, and Csd3 from *H. pylori* 26695 strain [Csd1 (HP_Csd1; SWISS-PROT accession code O26068), Csd2 (HP_Csd2; O26069), and Csd3 (HP_Csd3; O25247)], *S. aureus* LytM (SA_LytM; O33599), and *S. simulans* lysostaphin (SS_LytM; P10547). Tyr204 of *S. aureus* LytM is marked by a red star. Conserved residues of the characteristic motifs are colored in blue.

2.3.7. Discussion

Over the past years, many members of the M23 metallopeptidase family have been identified and biochemically characterized among the Gram-positive and Gram-negative bacteria. They have been implicated in a variety of important processes, including cell division, cell elongation, cell-shape determination, and sporulation (Uehara et al., 2009; Sycuro et al., 2010; Goley et al., 2010; Möll et al., 2010; Poggio et al., 2010; Bonis et al., 2010; Meisner and Moran, 2011; Singh et al., 2012). Many of its members are recognized by the catalytic LytM domain, which possesses two characteristic motifs for Zn^{2+}-binding (HxxxD and HxH) and a conserved histidine residue for catalysis. Some members of the M23 metallopeptidase family are expected to be non-peptidases, as their LytM domains have the degenerated active site. There are many examples for inactive homologs of enzymes acquiring new functions as binding proteins (Pils and Schultz, 2004). The *E. coli* LytM proteins (EnvC and NlpD) are involved in the cell division process by activating the septal peptidoglycan hydrolysis by amidase (Uehara et al., 2010). The LytM domain of EnvC lacks all five key residues and its structure showed that the catalytic metal ion is missing from the active site. Mutational analyses revealed that residues around the

degenerated active site are critical for amidase activation in vivo and in vitro (Peters et al., 2013). In *B. subtilis*, the membrane protein SpoIIQ functions as a structural component by interacting with another membrane protein SpoIIIAH to form the core of a channel that connects the two compartments during sporulation (Meisner and Moran, 2011; Levdikov et al., 2012; Meisner et al., 2012). SpoIIQ contains an extracellular LytM domain having the degenerated active site that misses one of three metal-binding residues (Ser119 instead of His of the HxxxD motif) and the proposed catalytic residue (Ser169 instead of His), resulting in the absence of a metal ion. The structure of the SpoIIQ and SpoIIIAH complex revealed that SpoIIIAH recognizes a region that protrudes from the N terminus of the SpoIIQ LytM domain. Both *H. pylori* Csd1 and Csd2 belong to the M23 metallopeptidase family. However, they differ significantly in their LytM domains; all five key residues are conserved in Csd1, whereas only three of the five key residues are conserved in Csd2.

Here, I have identified the 1:1 complex formation between Csd1 (HP1543) and Csd2 (HP1544) by solving the crystal structure of the $Csd1_{125–312}$-$Csd2_{121–308}$ heterodimer and also by performing SEC-MALS in solution. This finding is in agreement with the previous genetic study that demonstrated that the deletion of either *csd1* or *csd2*, or deletion of both

genes led to a similar increase in tetra-penta peptide cross-linked dimers in muropeptide composition and impaired helical twist resulting in the curved-rod morphology (Sycuro et al., 2010). Therefore, one can imagine that Csd1 and Csd2 might form a complex to function in *H. pylori* cells. Crystal structures reveal that the active site of the Csd1 LytM domain is bound with a Zn^{2+} ion but the Csd2 LytM domain is degenerate and no metal is bound. The present $Csd1_{125–312}$-$Csd2_{121–308}$ heterodimer is the first structure of the complex between two LytM homologs. The complex structure reveals that Csd1 and Csd2 make a heterodimer through their helical domains and dimerization does not affect the active site of the LytM domain. The helical domains of Csd1 and Csd2 are not conserved in other LytM proteins and show no structural similarity to other known protein structures. In both Csd1 and Csd2, the helical domain consists of a helix preceding the N-terminal side of the LytM domain and three helices following the C-terminal side of the LytM domain.

In *H. pylori*, the *ccmA* gene is adjacent to *csd1* gene in a three-gene locus (Sycuro et al., 2010). Deletion of the *ccmA* gene led to similar alterations in the muropeptide composition observed for single or double gene deletion for *csd1* and *csd2* (Sycuro et al., 2010). CcmA lacks any recognizable peptidase motif but is a homolog of bactofilins, which are a

widespread class of bacterial filament-forming proteins and serve as cytoskeletal scaffolds in various cellular pathways (Vasa et al., 2015). The highly polymerized CcmA protein may interact with Csd1 and Csd2 proteins to facilitate their localization in forming potential cellular machinery for precisely processing peptidoglycan. However, I could not test the interactions, because the CcmA protein formed high molecular aggregates when it was overexpressed in *E. coli*, as one might anticipate from its inherent property to form filaments.

A recent study described the synthesis of a phosphinic acid-based inhibitor against Csd4 from *H. pylori* (Liu et al. 2016). It demonstrated that the hydrophilic small molecule inhibitor of Csd4 can cross the outer membrane of *H. pylori* and cause cell straightening, suggesting that Csd4 is a novel target for antibiotic development. Similarly, small molecule inhibitors of Csd1 could also be developed for the discovery of potential new antibiotics. My work lays the foundation for such efforts.

2.4. References

Afonine, P. V., Moriarty, N. W., Mustyakimov, M., Sobolev, O. V., Terwilliger, T. C., Turk, D., Urzhumtsev, A., and Adams, P. D. (2015) FEM: feature-enhanced map. Acta Crystallogr. sect. D Biol. Crystallogr. **71**:646–666.

An, D. R., Kim, H. S., Kim, J., Im, H. N., Yoon, H. J., Yoon, J. Y., Hesek, D., Lee, M., Mobashery, S., Kim, S. J., Lee, B. I., and Suh, S. W. (2015) Structure of Csd3 from *Helicobacter pylori*, a cell shape-determining metallopeptidase. Acta Crystallogr. sect. D Biol. Crystallogr. **71**:675–686.

Berg, H. C., and Turner, L. (1979) Movement of microorganisms in viscous environments. Nature **278**:349–351.

Brünger, A. T. (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature **355**:472–474.

Bonis, M., Ecobichon, C., Guadagnini, S., Prévost, M. C., and Boneca, I. G. (2010) A M23B family metallopeptidase of *Helicobacter pylori* required for cell shape, pole formation and virulence. Mol. Microbiol. **78**:809–819.

Chan, A. C., Blair, K. M., Liu, Y., Frirdich, E., Gaynor, E. C., Tanner, M. E., Salama, N. R., and Murphy, M. E. (2015) Helical shape of *Helicobacter*

pylori requires an atypical glutamine as a zinc ligand in the carboxypeptidase Csd4. J. Biol. Chem. **290**:3622–3638.

Chaput, C., Ecobichon, C., Cayet, N., Girardin, S. E., Werts, C., Guadagnini, S., Prévost, M. C., Mengin-Lecreulx, D., Labigne, A., and Boneca, I. G. (2006) Role of AmiA in the morphological transition of *Helicobacter pylori* and in immune escape. PLoS Pathog. **2**:e97.

Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. sect. D Biol. Crystallogr. **66**:12–21.

Costa, K., Bacher, G., Allmaier, G., Dominguez-Bello, M. G., Engstrand, L., Falk, P., de Pedro, M. A., and García-del Portillo, F. (1999) The morphological transition of *Helicobacter pylori* cells from spiral to coccoid is preceded by a substantial modification of the cell wall. J. Bacteriol. **181**:3710–3715.

Diederichs, K., and Karplus, P. A., (1997) Improved R-factors for diffraction data analysis in macromolecular crystallography. Nat. Struct. Biol. **4:**269–275.

Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. Acta Crystallogr. sect. D Biol. Crystallogr. **66**:486–501.

Franceschi, F., Gasbarrini, A., Polyzos, S. A., and Kountouras, J. (2015) Extragastric diseases and *Helicobacter pylori*. Helicobacter **20** (Suppl. S1):40–46.

Frirdich, E., Vermeulen, J., Biboy, J., Soares, F., Taveirne, M. E., Johnson, J. G., DiRita, V. J., Girardin, S. E., Vollmer, W., and Gaynor, E. C. (2014) Peptidoglycan LD-carboxypeptidase Pgp2 influences *Campylobacter jejuni* helical cell shape and pathogenic properties and provides the substrate for the DL-carboxypeptidase Pgp1. J. Biol. Chem. **289**:8007–8018.

Goley, E. D., Comolli, L. R., Fero, K. E., Downing, K. H., and Shapiro, L. (2010) DipM links peptidoglycan remodeling to outer membrane organization in *Caulobacter*. Mol. Microbiol. **77**:56–73.

Gouet, P., Robert, X., and Courcelle, E. (2003) Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. **31:**3320–3323.

Grabowska, M., Jagielska, E., Czapinska, H., Bochtler, M., and Sabala, I. (2015). High resolution structure of an M23 peptidase with a substrate analogue. Sci. Rep. **5**:14833.

Hazell, S. L., Lee, A., Brady, L., and Hennessy, W. (1986) *Campylobacter pyloridis* and *gastritis*: association with intercellular spaces and adaptation to an environment of mucus as important factors in colonization of the gastric epithelium. J. Infect. Dis. **153**:658–663.

van Heijenoort, J. (2011) Peptidoglycan hydrolases of *Escherichia coli*. Microbiol. Mol. Biol. Rev. **75**:636–663.

Heinig, M., and Frishman, D. (2004) STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. Nucleic Acids Res. **32:**W500–W502.

Kim, H. S., Kim, J., Im, H. N., An, D. R., Lee, M., Hesek, D., Mobashery, S., Kim, J. Y., Cho, K., Yoon, H. J., Han, B. W., Lee, B. I., and Suh, S. W. (2014). Structural basis for the recognition of muramyltripeptide by *Helicobacter pylori* Csd4, a D,L-carboxypeptidase controlling the helical cell shape. Acta Crystallogr. sect. D Biol. Crystallogr. **70**:2800–2812.

Kim, H. S., Im, H. N., An, D. R., Yoon, J. Y., Jang, J. Y., Mobashery, S., Hesek, D., Lee, M., Yoo, J., Cui, M., Choi, S., Kim, C., Lee, N. K., Kim, S. J., Kim, J. Y., Bang, G., Han, B. W., Lee, B. I., Yoon, H. J., and Suh, S. W. (2015) The cell shape-determining Csd6 protein from Helicobacter pylori constitutes a new family of L,D-carboxypeptidase. J. Biol. Chem. **290**:25103–25117.

Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. **372**:774–797.

Lee, M., Hesek, D., Blázquez, B., Lastochkin, E., Boggess, B., Fisher, J. F., Mobashery, S. (2015) Catalytic spectrum of the penicillin-binding protein 4 of *Pseudomonas aeruginosa*, a nexus for the induction of β-lactam antibiotic resistance. J. Am. Chem. Soc. **137**: 190–200.

Lee, M., Hesek, D., Llarrull, L. I., Lastochkin, E., Pi, H., Boggess, B., and Mobashery, S. (2013) Reactions of all *Escherichia coli* lytic transglycosylases with bacterial cell wall. J. Am. Chem. Soc. **135:** 3311–3314.

Lee, T. K., and Huang, K.C. (2013) The role of hydrolases in bacterial cell-wall growth. Curr. Opin. Microbiol. **16**:760–766.

Lertsethtakarn, P., Ottemann, K. M., and Hendrixson, D. R. (2011) Motility and chemotaxis in *Campylobacter* and *Helicobacter*. Annu. Rev. Microbiol. **65**:389–410.

Levdikov, V. M., Blagova, E. V., McFeat, A., Fogg, M. J., Wilson, K. S., and Wilkinson, A. J. (2012) Structure of components of an intercellular channel complex in sporulating *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA **109**:5441–5445.

Liu, Y., Frirdich, E., Taylor, J. A., Chan, A. C., Blair K. M., Vermeulen, J., Ha, R., Murphy, M. E., Salama, N. R., Gaynor, E. C., and Tanner, M. E. (2016) A bacterial cell shape-determining inhibitor. ACS Chem. Biol. (DOI: 10.1021/acschembio.5b01039) [Epub ahead of print]

Malfertheiner, P., Selgrad, M., and Bornschein, J. (2012) *Helicobacter pylori*: clinical management. Curr. Opin. Gastroenterol. **28**:608–614.

Meisner, J., Maehigashi, T., André, I., Dunham, C. M., and Moran, C. P. (2012) Structure of the basal components of a bacterial transporter. Proc. Natl. Acad. Sci. USA **109**:5446–5451.

Meisner, J., and Moran, C. P. (2011) A LytM domain dictates the localization of proteins to the mother cell-forespore interface during bacterial endospore formation. J. Bacteriol. **193**:591–598.

Meroueh, S. O., Bencze, K. Z., Hesek, D., Lee, M., Fisher, J. F., Stemmler, T. L., and Mobashery, S. (2006) Three-dimensional structure of the bacterial cell wall peptidoglycan. Proc. Natl. Acad. Sci. USA **103**:4404–4409.

Möll, A., Schlimpert, S., Briegel, A., Jensen, G. J., and Thanbichler, M. (2010) DipM, a new factor required for peptidoglycan remodeling during cell division in Caulobacter crescentus. Mol. Microbiol. **77**:90–107.

Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. sect. D Biol. Crystallogr. **53**:240–255.

Ottemann, K. M., and Lowenthal, A. C. (2002) *Helicobacter pylori* uses motility for initial colonization and to attain robust infection. Infect. Immun. **70**:1984–1990.

Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in osillation mode. Methods Enzymol. **276**:307–326.

Peters, N. T., Morlot, C., Yang, D. C., Uehara, T., Vernet, T., and Bernhardt, T. G. (2013) Structure-function analysis of the LytM domain of EnvC, an activator of cell wall remodelling at the *Escherichia coli* division site. Mol. Microbiol. **89**:690–701.

Pils, B., and Schultz, J. (2004) Inactive enzyme-homologues find new function in regulatory processes. J. Mol. Biol. **340**:399–404.

Poggio, S., Takacs, C. N., Vollmer, W., and Jacobs-Wagner, C. (2010) A protein critical for cell constriction in the Gram-negative bacterium Caulobacter crescentus localizes at the division site through its peptidoglycan-binding LysM domains. Mol. Microbiol. **77**:74–89.

Read, R. J., and Schierbeek, A. J. (1988) A phased translation function. J. Appl. Crystallogr. **21:**490–495.

Roesler, B. M., Rabelo-Gonçalves, E. M., and Zeitune, J. M. (2014) Virulence factors of *Helicobacter pylori*: A review. Clin. Med. Insights Gastroenterol. **7**:9–17.

Sabala, I., Jagielska, E., Bardelang, P. T., Czapinska, H., Dahmn, S. O., Sharpe, J. A., James R., Than, M. E., Thomas, N. R., and Bochtler, M. (2014) Crystal structure of the antimicrobial peptidase lysostaphin from *Staphylococcus simulans*. FEBS J. **281**:4112–4122.

Scheffers, D. J., and Pinho, M. G. (2005) Bacterial cell wall synthesis: new insights from localization studies. Microbiol. Mol. Biol. Rev. **69:**585–607.

Schreiber, S., Konradt, M., Groll, C., Scheid, P., Hanauer, G., Werling, H. O., Josenhans, C., and Suerbaum, S. (2004) The spatial orientation of *Helicobacter pylori* in the gastric mucus. Proc. Natl. Acad. Sci. USA **101**:5024–5029.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., and Higgins, D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. **7**:539.

Singh, S. K., SaiSree, L., Amrutha, R. N., and Reddy, M. (2012) Three redundant murein endopeptidases catalyse an essential cleavage step in

peptidoglycan synthesis of *Escherichia coli* K12. Mol. Microbiol. **86**:1036–1051.

Spencer, J., Murphy, L. M., Conners, R., Sessions, R. B., and Gamblin, S. J. (2010) Crystal structure of the LasA virulence factor from *Pseudomonas aeruginosa*: substrate specificity and mechanism of M23 metallopeptidases. J. Mol. Biol. **396**:908–923.

Sycuro, L. K., Pincus, Z., Gutierrez, K. D., Biboy, J., Stern, C. A., Vollmer, W., and Salama, N. R. (2010) Peptidoglycan crosslinking relaxation promotes *Helicobacter pylori*'s helical shape and stomach colonization. Cell **141:**822–833.

Sycuro, L. K., Rule, C. S., Petersen, T. W., Wyckoff, T. J., Sessler, T., Nagarkar, D. B., Khalid, F., Pincus, Z., Biboy, J., Vollmer, W., and Salama, N. R. (2013) Flow cytometry-based enrichment for cell shape mutants identifies multiple genes that influence *Helicobacter pylori* morphology. Mol. Microbiol. **90**:869–883.

Sycuro, L. K., Wyckoff, T. J., Biboy, J., Born, P., Pincus, Z., Vollmer, W., and Salama, N. R. (2012) Multiple peptidoglycan modification networks modulate *Helicobacter pylori*'s cell shape, motility, and colonization potential. PLoS Path. **8**:e1002603.

Uehara, T., Dinh, T., and Bernhardt, T. G. (2009) LytM-domain factors are required for daughter cell separation and rapid ampicillin-induced lysis in *Escherichia coli*. J. Bacteriol. **191**:5094–5107.

Uehara, T., Parzych, K. R., Dinh, T., and Bernhardt, T. G. (2010) Daughter cell separation is controlled by cytokinetic ring-activated cell wall hydrolysis. EMBO J. **29**:1412–1422.

Vagin, A., and Teplyakov, A. (2010) Molecular replacement with MOLREP. Acta Crystallogr. sect. D Biol. Crystallogr. **66**:22–25.

Vasa, S., Lin, L., Shi, C., Habenstein, B., Riedel, D., Kühn, J., Thanbichler, M., and Lange, A. (2015) β-Helical architecture of cytoskeletal bactofilin filaments revealed by solid-state NMR. Proc. Natl. Acad. Sci. USA **112**:E127– E136.

Vollmer, W. (2012) Bacterial growth does require peptidoglycan hydrolases. Mol. Microbiol. **86:**1031–1035.

Vollmer, W., and Bertsche, U. (2008) Murein (peptidoglycan) structure, architecture and biosynthesis in *Escherichia coli*. Biochim. Biophys. Acta **1778:**1714–1734.

Vollmer, W., Blanot, D., and De Pedro, M. A. (2008a) Peptidoglycan structure and architecture. FEMS Microbiol. Rev. **32:**149–167.

Vollmer, W., Joris, B., Charlier, P., and Foster, S. (2008b) Bacterial peptidoglycan (murein) hydrolases. FEMS Microbiol. Rev. **32:**259–286.

Weidel, W., and Pelzer, H. (1964) Bag-shaped macromolecules: A new outlook on bacterial cell walls. Adv. Enzymol. Relat. Areas Mol. Biol. **26**:193–232.

Weiss, M. S. (2001) Global indicators of X-ray data quality. J. Appl. Cryst. **34:**130–135.

Worku, M. L., Sidebotham, R. L., Walker, M. M., Keshavarz, T., and Karim, Q. N. (1999) The relationship between *Helicobacter pylori* motility, morphology and phase of growth: implications for gastric colonization and pathology. Microbiology **145**:2803–2811.

Wyckoff, T. J., Taylor, J. A., and Salama, N. R. (2012) Beyond growth: novel functions for bacterial cell wall hydrolases. Trends Microbiol. **20**:540–547.

국문초록

Structural Studies of Three Cell Shape Determining Proteins from *Helicobacter pylori*: Csd1, Csd2, and Csd3

헬리코박터 파일로리균 유래의 세 가지 세포 모양 결정인자 단백질 (Csd1, Csd2 와 Csd3)의 구조연구

헬리코박터 파일로리균은 위암을 포함한 다양한 위장관 질병을 유발한다. 이들의 인체 내 감염은 위 점막에서의 효과적인 운동성을 필요로 하며, 이는 헬리코박터균의 나선형 세포 모양에 의해 결정된다. 이러한 헬리코박터균의 특징적인 세포 모양은 세포 모양 결정인자 (Csd) 단백질들에 의해 결정되며, 이들은 대부분 펩티도글라이칸의 교차 결합을 와해하거나 줄기 펩타이드(stem peptide)의 펩타이드 결합을 절단함으로써 세포 모양을 결정하는 중요한 역할을 한다. 세 개의 Csd 단백질 (Csd1, Csd2, Csd3)은 동일한 M23B metallopeptidase 패밀리에 속하며, 펩티도글라이칸 dimer muropeptide 의 교차 결합을 끊는 D,D-endopeptidase 로 작용한다고 알려졌다. 이들 중, Csd3 는 D,D-endopeptidase 뿐만 아니라 muramyl pentapeptide 의 마지막 펩타이드 결합을 끊는 D,D-carboxypeptidase 로도 기능하는 이중 활성 단백질이다. 본 연구에서는 헬리코박터균에서 Csd 단백질의 역할에 대한 이해를 증진시키기 위해 세 개의 Csd 단백질, 즉 HP1543 (*H. pylori* Csd1), HP1544 (*H. pylori* Csd2) 과 HP0506 (*H. pylori* Csd3)에 대한 삼차원 구조를 규명하였다. *H. pylori* Csd3 는 단량체 상태로 존재하며, 카르복시 말단의 LytM domain 을 포함한 세 개의 도메인으로 구성되어 있다. Csd3 LytM domain 은 한 개의 Zn^{2+} 이온을 포함한 활성 부위를 포함하고 있으며, 이는 다른 M23

metallopeptidase 단백질들의 활성 도메인과 매우 유사하다. 하지만, Csd3 의 활성 부위는 아미노 말단의 도메인에 의해 막혀있는 비활성 상태이다. 독자적으로 발현된 *H. pylori* Csd2 단백질은 단량체와 이합체 평형 상태로 존재하며, *H. pylori* Csd1 와의 안정적인 이종이합체를 형성한다. Csd1 과 Csd2 단량체의 구조는 서로 상당히 유사하며, 이들은 helical domain 과 LytM domain 으로 구성되어있다. 두 단백질의 helical domain 은 이합체를 형성하는데 중요한 역할을 한다. 두 단백질의 LytM domain 은 전체 구조가 상당히 유사하지만, 각자의 활성 부위에 큰 차이점을 보인다. Csd1 LytM domain 은 잘 보존된 세 개의 아미노산 리간드와 두 개의 물 분자를 통하여 결합한 Zn^{2+} 이온을 포함한 활성 부위를 가지고 있는 반면, Csd2 LytM domain 은 불완전한 리간드와 금속 이온이 포함되지 않은 활성 부위를 갖는다. 세 개의 Csd 단백질은 헬리코박터균의 LytM 동족체이지만, 각 단백질의 LytM domain 이 구조적인 특이성을 보인다. 본인의 Csd 단백질에 대한 구조 연구는 헬리코박터균내의 세포벽 조절 과정을 이해하는데 필요한 중요한 정보를 제공한다.

주요어: *Helicobacter pylori* / cell shape determinant / Csd1 / Csd2 / Csd3 / HP1543 / HP1544 / HP0506 / peptidoglycan / D,D-endopeptidase / D,D-carboxypeptide / M23B family metallopeptidase / LytM domain